



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 131014

TO: Terra Gibbs
Location: REM/2D10/2C18
Art Unit: 1635
Monday, August 30, 2004

Case Serial Number: 10/029115

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

Search Notes

RUSH

THIS PAGE BLANK (USPTO)

Gibbs, Terra

From: Gibbs, Terra
Sent: Thursday, August 19, 2004 5:21 PM
To: STIC-Biotech/ChemLib
Subject: RE: Sequence comparison

Please modify the search below. The claims are drawn to a nucleic acid at least 98% or 95% identical to SEQ ID NO:1 of USSN 10/029,115.

I need the parameters tweaked such that the sequences are at least 98% identical. If they aren't at least 98% identical, please tweak the parameters such that the sequences are at least 95% identical.

Thank You!

-----Original Message-----

From: Gibbs, Terra
Sent: Thursday, August 19, 2004 5:18 PM
To: STIC-Biotech/ChemLib
Subject: Sequence comparison

1. Could you please compare SEQ ID NO:11 of US 6,656,716 with SEQ ID NO:1 of USSN 10/029,115? I need the parameters tweaked such that the sequences are at least 95% identical.
2. Could you please compare Accession number AB035698 with SEQ ID NO:1 of USSN 10/029,115? I need the parameters tweaked such that the sequences are at least 95% identical.

*Terra Cotta Gibbs, Ph.D.
Art Unit 1635
Remsen Building 2D10
Mailbox 2C18
571-272-0758*

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 28, 2004, 00:43:54, Search time 29 Seconds
(without alignments)
2355.726 Million cell updates/sec

Title: US-10-029-115-2

Perfect score: 6929
Sequence: 1 MGDPAAPASLDIDSLALRD.....SCGSSQVYFNTLRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2002000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	99.7	1332	1	MAK6_HUMAN
2	6464	93.3	1308	1	MAK6_MOUSE
3	4336	62.6	1360	1	TNIK_HUMAN
4	4017.5	58.0	1239	1	MAK4_HUMAN
5	3951.5	57.0	1233	1	MAK4_MOUSE
6	3286	47.4	916	1	TNIK_MOUSE
7	2646	38.2	1080	1	MIL5_MOUSE
8	728	10.7	894	1	MAK3_HUMAN
9	715.5	10.3	862	1	MAK3_RAT
10	693.5	10.0	491	1	STK3_HUMAN
11	693.5	10.0	669	1	HPO_DROME
12	691	10.0	847	1	MAK5_MOUSE
13	687	9.9	846	1	MAK5_HUMAN
14	681.5	9.8	426	1	STK5_HUMAN
15	675.5	9.7	426	1	STK5_MOUSE
16	667	9.6	1080	1	NPK1_YEAST
17	660.5	9.5	833	1	MAK1_HUMAN
18	656.5	9.5	443	1	STK4_HUMAN
19	653.5	9.4	819	1	MAK2_HUMAN
20	644	9.3	821	1	STK4_MOUSE
21	638	9.2	966	1	MAK2_MOUSE
22	636.5	9.1	691	1	MAK3_MOUSE
23	630.5	9.1	827	1	NINQ_DROME
24	629	9.0	471	1	STUD_YEAST
25	620.5	8.0	982	1	STUD_MOUSE
26	552	8.0	490	1	SPK1_YEAST
27	551	7.9	658	1	CLAK_YEAST
28	548	7.7	971	1	CLAK_MOUSE
29	546	7.3	652	1	NPK1_MOUSE
30	533.5	7.3	556	1	SPAK_MOUSE
31	507	7.3			
32	505	7.3			
33	504	7.3			

ALIGNMENTS

RESULT 1	MAK6_HUMAN	STANDARD:	PRT: 1332 AA.
ID	MAK6_HUMAN	Q9N4C8; Q9P1X1; Q9P2R8;	
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Mitogen-activated protein kinase kinase kinase kinase 6 (MEKKK 6)		
DE	(MAPK/ERK kinase kinase kinase 6) (GCK family kinase MINK).		
DE	(MAPK/ERK kinase kinase kinase 6) (GCK family kinase MINK).		
GN	MAPK6 OR MINK.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Brain;		
RX	MEDLINE=20175403; PubMed=10708748;		
RA	Dan I., Watanabe N.M., Kobayashi T., Yamashita-Suzuki K., Fukagaya Y.,		
RA	Kajikawa E., Kimura W.K., Nakashima T.M., Matsumoto K.,		
RT	Ninomiya-Tsujii J., Kusumi A.;		
RT	Molecular cloning of MINK, a novel member of mammalian GCK family		
RT	kinases, which is up-regulated during postnatal mouse cerebral		
RT	development.;		
RT	FEBS Lett. 469:19-23 (2000).		
RN	[2]		
RN	SEQUENCE FROM N.A. (ISOFORM 3).		
RC	TISSUE=Lymph;		
RX	MEDLINE=22386257; PubMed=12477932;		
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Matulis M., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Caavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Cantin P., Prange C.,		
RA	Rah S.S., Loguailano N.A., Peters G.J., Abrahamson R.D., Mullaly S.C.,		
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gurrane P.H.,		
RA	Richards S., Wooley K.C., Hale S., Garcia A.H., Gay L.J., Huix S.W.,		
RA	Villalón D.X., Huix D.H., Sodergren E.O., Lu X., Gibbs R.A.,		
RA	Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Rutledge Y.S.N., Krzywicki M.I., Skalska V., Smalins D.E.,		
RA	Schneer A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences.;		
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
CC	- FUNCTION: serine/threonine kinase that may play a role in the		
CC	response to environmental stress. Appears to act upstream of the		
CC	c-Jun N-terminal pathway (by similarity).		
CC	- FUNCTION: May play a role in the development of the brain (by		
CC	similarity).		
CC	- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a		

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Mon Aug 30 09:54:26 2004

us-10-02

Db 1201 YIPVHIQSQITPHAIIFLPNTDGMEMLLCYEDEGVYVNTYGRITIKDVVLOWGEMPTSVAY 1260
QY 1241 ICSNQIMGWGEKAIEIRSVETGHLDCGVFMKRAQLKFLCERNDKVFFASVRSGGSSQVY 1300
Db 1261 ICSNQIMGWGEKAIEIRSVETGHLDCGVFMKRAQLKFLCERNDKVFFASVRSGGSSQVY 1320
QY 1301 FMTLNRNCIMNW 1312
Db 1321 FMTLNRNCIMNW 1332

THIS PAGE BLANK (uspto)

Db 241 AACATCGCCACCTACTACGAGCCTTCATCAAGAGAGAGCCGCCCGGGAACGATGACCG 300
Qy 307 CTCTGGCTGGTGATGGAGTTCTGTGGTGTCTGTGGTGTGAGTGAAGTGAAGAACACA 366
Db 301 CTCTGGCTGGTGATGGAGTTCTGTGGTGTCTGTGGTGTGAGTGAAGTGAAGAACACA 360
Qy 367 AAAGGCAACGCCCTGAAGAGAGACTGTATCGGCTATATCTGAGGAGAGATCTCAGGGGT 426
Db 361 AAAGGCAACGCCCTGAAGAGAGACTGTATCGGCTATATCTGAGGAGAGATCTCAGGGGT 420
Qy 427 CTGGCCCATCTCCATGCCACCAAGGTGATCCATCGAGACATCAAGGGGCGAGAATGTCTG 486
Db 421 CTGGCCCATCTCCATGCCACCAAGGTGATCCATCGAGACATCAAGGGGCGAGAATGTCTG 480
Qy 487 CTGACAGAGAATGCTGAGGTCAAGCTAGTGGATTTTGGGGTGAAGTCTCAGCTGACCCGC 546
Db 481 CTGACAGAGAATGCTGAGGTCAAGCTAGTGGATTTTGGGGTGAAGTCTCAGCTGACCCGC 540
Qy 547 ACGTGGGCGAGGAGAACATTTTCATTGGGACTCCCTACTGATGGCTCCAGAGTTCATC 606
Db 541 ACGTGGGCGAGGAGAACATTTTCATTGGGACTCCCTACTGATGGCTCCAGAGTTCATC 600
Qy 607 GCCTGTGATGAGAACCCCTGATGCCACCTATGATTAACAGGAGTGATATTTGGTCTCTAGA 666
Db 601 GCCTGTGATGAGAACCCCTGATGCCACCTATGATTAACAGGAGTGATATTTGGTCTCTAGA 660
Qy 667 ATCACAGCCATCGAGATGCGAGGAGGCCCCCTCTGTGTGATGACACCCCATGCGA 726
Db 661 ATCACAGCCATCGAGATGCGAGGAGGCCCCCTCTGTGTGATGACACCCCATGCGA 720
Qy 727 GCCCTCTTCTCATTTCTCGGAACCTCCGCCAGGCTCAAGTCCAGAGTGGTCTAAG 786
Db 721 GCCCTCTTCTCATTTCTCGGAACCTCCGCCAGGCTCAAGTCCAGAGTGGTCTAAG 780
Qy 787 AAGTTCAATTGACTTCACTGACATGTCTCATCAAGACTTTACTGAGCGGCCACCCACG 846
Db 781 AAGTTCAATTGACTTCACTGACATGTCTCATCAAGACTTTACTGAGCGGCCACCCACG 840
Qy 847 GAGCAGTACTGAAGTTTCCCTTCAATCGGGAACAGCCCAACGAGCGGAGTCCGCAATC 906
Db 841 GAGCAGTACTGAAGTTTCCCTTCAATCGGGAACAGCCCAACGAGCGGAGTCCGCAATC 900
Qy 907 CAGCTTAAGGACCACTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
Db 901 CAGCTTAAGGACCACTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 967 TATGAGTACAGCGGCGAGGAGGAGAGAGATGACAGCCATGAGAGAGAGAGAGAGAGAGAG 1026
Db 961 TATGAGTACAGCGGCGAGGAGGAGAGAGATGACAGCCATGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1027 TCCATCATGAACGTGTGCTGAGAGTGCATCTTACGCGGGAGTTTCTCGGCTCCAGCAG 1086
Db 1021 TCCATCATGAACGTGTGCTGAGAGTGCATCTTACGCGGGAGTTTCTCGGCTCCAGCAG 1080
Qy 1087 GAATATAGAGCACTCAGAGGCTTTTAAACAGCAGCAGCAGCAGCTGCAGCAGCAGCAGCAG 1146
Db 1081 GAATATAGAGCACTCAGAGGCTTTTAAACAGCAGCAGCAGCAGCTGCAGCAGCAGCAGCAG 1140
Qy 1147 CGAGACCCCGGCGACACATCAAAACCTGTGACACCGCAGCGGCGAGCGGCGATAGAGAG 1206
Db 1141 CGAGACCCCGGCGACACATCAAAACCTGTGACACCGCAGCGGCGAGCGGCGATAGAGAG 1200
Qy 1207 CAGAGAGGAGAGCGGCGCGCTGTGAGAGAGCAACAGCGGCGGGAACGCGGAGCAGCGGAAG 1266
Db 1201 CAGAGAGGAGAGCGGCGCGCTGTGAGAGAGCAACAGCGGCGGGAACGCGGAGCAGCGGAAG 1260
Qy 1267 CTGCAGGAGAGAGCGGCGCGCTGTGAGAGAGCAACAGCGGCGGGAACGCGGAGCAGCGGAAG 1326
Db 1261 CTGCAGGAGAGAGCGGCGCGCTGTGAGAGAGCAACAGCGGCGGGAACGCGGAGCAGCGGAAG 1320
Qy 1327 GAGCGGCGGAGCGGAGCGGAG 1386
Db 1321 GAGCGGCGGAGCGGAGCGGAG 1380

Qy 1387 CAGTCAGAACGTCTTCCAGAGGCGAGCTGCAGCAGAGCATGCTACCTCAAGTCTCCCTGCGAG 1446
Db 1381 CAGTCAGAACGTCTTCCAGAGGCGAGCTGCAGCAGAGCATGCTACCTCAAGTCTCCCTGCGAG 1440
Qy 1447 CAGCAGCAACAGCAGCAGCAGCTTCAGAAACAGCAGCAGCAGCAGCTTCCTGCTGGGGGAC 1506
Db 1441 CAGCAGCAACAGCAGCAGCAGCTTCAGAAACAGCAGCAGCAGCAGCTTCCTGCTGGGGGAC 1500
Qy 1507 AGGAAGCCCTGTATACCATTTATGGTCGGGGCATGAATCCCGCTGACAAACAGACCTTGCGCC 1566
Db 1501 AGGAAGCCCTGTATACCATTTATGGTCGGGGCATGAATCCCGCTGACAAACAGACCTTGCGCC 1560
Qy 1567 CGAGAGGTAGAAGAGAGAAACAGGATGAAACAGCAGCAGCAAACTCTCTCTTGGCCCAAGAGC 1626
Db 1561 CGAGAGGTAGAAGAGAGAAACAGGATGAAACAGCAGCAGCAAACTCTCTCTTGGCCCAAGAGC 1620
Qy 1627 AAGCCAGCAGCAGCAGCGGCGCTGAGCCGCCCATCCCGCAGGCTCCCGAGGGCCCCCAGGA 1686
Db 1621 AAGCCAGCAGCAGCAGCGGCGCTGAGCCGCCCATCCCGCAGGCTCCCGAGGGCCCCCAGGA 1680
Qy 1687 CCCCTTTTCCAGACTCTCTCTATGACAGGCGCGGTGGAGCCCCCAGGAGGGAACCGACACA -- 1744
Db 1681 CCCCTTTTCCAGACTCTCTCTATGACAGGCGCGGTGGAGCCCCCAGGAGGGAACCGACACAAG 1740
Qy 1745 -----AG 1746
Db 1741 AGCCTGGTGGCACACCGGGTCCCACTGAAGCCATATATGACGACCTGTACCCCGATCCCGAG 1800
Qy 1747 TCCCTGAGAGACAGACCCAGCCAGAAACCTGTGCTGCTTCCAGGCTCCCATGACCCCGAC 1806
Db 1801 TCCCTGAGAGACAGACCCAGCCAGAAACCTGTGCTGCTTCCAGGCTCCCATGACCCCGAC 1860
Qy 1807 CCTGCCATCCCGCGACCACTGTGCCACGCCAGTGTGCCGAGGAGCTGTATCCTCGCCAGAAT 1866
Db 1861 CCTGCCATCCCGCGACCACTGTGCCACGCCAGTGTGCCGAGGAGCTGTATCCTCGCCAGAAT 1920
Qy 1867 TCAGACCCCGAGCTTGAAGACCTTGGGCCCGACGCCGGAATCCCGCAGGCTTGGGTGCGGCCA 1926
Db 1921 TCAGACCCCGAGCTTGAAGACCTTGGGCCCGACGCCGGAATCCCGCAGGCTTGGGTGCGGCCA 1980
Qy 1927 GATAAACAGGCGCCCAACAGGTGCTCAGAGGACCTCATCTATGCGACCTGTCCCTTTAAC 1986
Db 1981 GATAAACAGGCGCCCAACAGGTGCTCAGAGGACCTCATCTATGCGACCTGTCCCTTTAAC 2040
Qy 1987 ACCAGTGGGCGGAGGCTTCCCGCCAGCCAGCAGTGTGCGCAGACCTTCGACGACAAAC 2046
Db 2041 ACCAGTGGGCGGAGGCTTCCCGCCAGCCAGCAGTGTGCGCAGTGTGCGC ----- 2085
Qy 2047 TCCGCTTGGCAATCTATCTGCAAGGCGGGGAGAGCGGGGACCCCAAGAGCTTCAGGG 2106
Db 2086 ----- 2085
Qy 2107 CCCCTGTCTAGCCCCCTGGCGCCGCCCAACGCTCTAGTAACCCCGACCTCAGAGGAGC 2166
Db 2086 -----AGTAACCCCGACCTCAGAGGAGC 2109
Qy 2167 GACCTTGGTGGGAAACGCTCGGACAGCTTCTTACGCTTCTACGCGGCACTTCCCGCAG 2226
Db 2110 GACCTTGGTGGGAAACGCTCGGACAGCTTCTTACGCTTCTACGCGGCACTTCCCGCAG 2169
Qy 2227 GCTGGCTCAGTGGAGCGGAACCGGTGGAGCTTCTTCAAACTGGGACAGCTTCCCTGTG 2286
Db 2170 GCTGGCTCAGTGGAGCGGAACCGGTGGAGCTTCTTCAAACTGGGACAGCTTCCCTGTG 2229
Qy 2287 CTCTCCCTTGGGAATAAGCCAGCCAGACCGCTCAGCGGCGAGGCGGCGCGCA 2346
Db 2230 CTCTCCCTTGGGAATAAGCCAGCCAGACCGCTCAGCGGCGAGGCGGCGCGCA 2289
Qy 2347 GACTTTGTGTGCTGAAGAGCGGACTCTGTGAGAGAGCGCTTGGCTTCCCAAGAGGCC 2406
Db 2290 GACTTTGTGTGCTGAAGAGCGGACTCTGTGAGAGAGCGCTTGGCTTCCCAAGAGGCC 2349


```
QY 2407 ATGGACTACTCGTCTCCAGCGAGGAGGTGGAAACAGTCGAGGACGACGAGGAGGAAGGC 2466
DB |||||
QY 2350 ATGGAATACTCGTCTGTCAGCGAGGAGGTGGAAACAGTCGAGGACGACGAGGAGGAAGGC 2409
DB |||||
QY 2467 GAAAGCGGGCCACGACGAGGAGGACGAGATACCCCTGGGGGCGCGAGCGGATGGGGATACA 2526
DB |||||
QY 2410 GAAAGCGGGCCACGACGAGGAGGACGAGATACCCCTGGGGGCGCGAGCGATGGGGATACA 2469
DB |||||
QY 2527 GACAGGGTCAGCACCAATGTTGTGTCCAGACGTCGAGAGAGATCACCGGAGCCAGGACCCCA 2586
DB |||||
QY 2470 GACAGCGTCAGCACCAATGTTGTGTCCAGACGTCGAGAGAGATCACCGGAGCCAGGACCCCA 2529
DB |||||
QY 2587 TACGGGGGCGGACCAATGTTGTGTCCAGGACACCCCTGAAGAGAGCGGAACTGCTGCAT 2646
DB |||||
QY 2530 TACGGGGGCGGACCAATGTTGTGTCCAGGACACCCCTGAAGAGAGCGGAACTGCTGCAT 2589
DB |||||
QY 2647 GCTGACAGCAATGGGGTACACAAACCTGCTGACGTGGTCCAGCCCAGGCCACTCACCCACC 2706
DB |||||
QY 2590 GCTGACAGCAATGGGGTACACAAACCTGCTGACGTGGTCCAGCCCAGGCCACTCACCCACC 2649
DB |||||
QY 2707 GAGAACAGCAAAAGGCCAAAGCCCAACCTCGAAGGATGGGAGTGGTACTACCACTCTCGT 2766
DB |||||
QY 2650 GAGAACAGCAAAAGGCCAAAGCCCAACCTCGAAGGATGGGAGTGGTACTACCACTCTCGT 2709
DB |||||
QY 2767 GGGCTGGTAAAGGCCCTCGGACAGAGCTCGTTACAGATGTTTGTGGATCTAGGGATCTAC 2826
DB |||||
QY 2710 GGGCTGGTAAAGGCCCTCGGACAGAGCTCGTTACAGATGTTTGTGGATCTAGGGATCTAC 2769
DB |||||
QY 2827 CAGCCTGAGGAGGATGGGACAGACATCCCCATCACAGCCCTAGTGGTGGAGGGGCACT 2886
DB |||||
QY 2770 CAGCCTGAGGAGGATGGGACAGACATCCCCATCACAGCCCTAGTGGTGGAGGGGCACT 2829
DB |||||
QY 2887 CGGCTCGACCACTGTCAGTACGACGTGAGGAAGGGTCTGTGCTCAACGTAATCCCAAC 2946
DB |||||
QY 2830 CGGCTCGACCACTGTCAGTACGACGTGAGGAAGGGTCTGTGCTCAACGTAATCCCAAC 2889
DB |||||
QY 2947 AACACCGGGCCACAGTCAGACCCCTGAGATCCCGAAGTACAGAAGCGATTCAACTCC 3006
DB |||||
QY 2890 AACACCGGGCCACAGTCAGACCCCTGAGATCCCGAAGTACAGAAGCGATTCAACTCC 2949
DB |||||
QY 3007 GAGATCCTCTGTGACGCCCTTTGGGGGGTCAACCTGCTGCTGGTGGGCAACCGGGCTG 3066
DB |||||
QY 2950 GAGATCCTCTGTGACGCCCTTTGGGGGGTCAACCTGCTGCTGGTGGGCAACCGGGCTG 3009
DB |||||
QY 3067 ATGTTGCTGGAACGAAAGTGGGACGAGGAGGTGTATGGAATCATTTGGGCGGACGCTTC 3126
DB |||||
QY 3010 ATGTTGCTGGAACGAAAGTGGGACGAGGCAAGGTGTATGGAATCATTTGGGCGGACGCTTC 3069
DB |||||
QY 3127 CAGCAGATGGATGCTGGAGGGGCTCAACCTGCTCATCACCATCTCAGGGAAAGGAAC 3186
DB |||||
QY 3070 CAGCAGATGGATGCTGGAGGGGCTCAACCTGCTCATCACCATCTCAGGGAAAGGAAC 3129
DB |||||
QY 3187 AAACCTGGGGTGTATTAACCTGCTCCTGGCTCCGGAAACAGATTTCTGCACAAATGACCCAGAA 3246
DB |||||
QY 3130 AAACCTGGGGTGTATTAACCTGCTCCTGGCTCCGGAAACAGATTTCTGCACAAATGACCCAGAA 3189
DB |||||
QY 3247 GTGGAGAAGAACGAGGGCTGGAACCAACGTTGGGGGGAATGAGAGGGCTGCGGGGCACTACCGT 3306
DB |||||
QY 3190 GTGGAGAAGAACGAGGGCTGGAACCAACGTTGGGGGGAATGAGAGGGCTGCGGGGCACTACCGT 3249
DB |||||
QY 3307 GTTGTGAAATACAGCGGATTAAGTTCTCGTGTATCGCCCTCAAGAGCTCCGTTGAGAGGTG 3366
DB |||||
QY 3250 GTTGTGAAATACAGCGGATTAAGTTCTCGTGTATCGCCCTCAAGAGCTCCGTTGAGAGGTG 3309
DB |||||
QY 3367 TATGCTGGGCCCCCAACCTACCAAAATTCATGGCCCTTCAAGTCTCTTTGCCGACCTC 3426
DB |||||
QY 3310 TATGCTGGGCCCCCAACCTACCAAAATTCATGGCCCTTCAAGTCTCTTTGCCGACCTC 3369
DB |||||
QY 3427 CCCACCGCCCTCTGTGTCGACCTTGAAGTGAAGAGGGGAGGGGCTCAAGTCAATC 3486
DB |||||
QY 3370 CCCACCGCCCTCTGTGTCGACCTTGAAGTGAAGAGGGGAGGGGCTCAAGTCAATC 3429
DB |||||
QY 3487 TATGGCTCCAGTGTGCTTCCATGCTGTGGATGTCGACTCGGGGAACAGCTATGACATC 3546
DB |||||
```

```
DB 3430 TATGGCTCCAGTGTGCTGCTTCCATGCTGTGGATGTTCGACTCGGGGAAACAGCTATGACATC 3489
QY |||||
DB 3547 TACATCCTGTGCATCCAGAGCCAGATCAGCGCCCATCATCATCTTCTCTCCCAAC 3606
QY |||||
DB 3490 TACATCCTGTGCATCCAGAGCCAGATCAGCGCCCATCATCATCTTCTCTCCCAAC 3549
QY |||||
QY 3607 ACCGACGGCATGAGATGCTGCTGTGTACGAGACGAGGGTGTCTACGTCAACACGTAC 3666
DB |||||
DB 3550 ACCGACGGCATGAGATGCTGCTGTGTACGAGACGAGGGTGTCTACGTCAACACGTAC 3609
QY |||||
QY 3667 GGGCGCATCATTAAGGATGTGCTGTGTGCTGAGTGGGGGAGATGCTTCTGTGGCTAC 3726
DB |||||
QY 3610 GGGCGCATCATTAAGGATGTGCTGTGTGCTGAGTGGGGGAGATGCTTCTGTGGCTAC 3669
QY |||||
QY 3727 ATCTGCTCCACACAGATAATGGGCTGGGGTGCAGAAAGCCATTGAGATCCGCTCTGTGGAG 3786
DB |||||
DB 3670 ATCTGCTCCACACAGATAATGGGCTGGGGTGCAGAAAGCCATTGAGATCCGCTCTGTGGAG 3729
QY |||||
QY 3787 ACGGGCCACCTCGACGGGGTCTTTCATGCAACAAACAGCTCAGAGGCTCAAGTCTCTGTGT 3846
DB |||||
DB 3730 ACGGGCCACCTCGACGGGGTCTTTCATGCAACAAACAGCTCAGAGGCTCAAGTCTCTGTGT 3789
QY |||||
QY 3847 GAGCGGAATGACAAAGGTGTTTTTTCCTCAGTCCGCTCTGGGGGAGGAGCAAGTTTAC 3906
DB |||||
QY 3790 GAGCGGAATGACAAAGGTGTTTTTTCCTCAGTCCGCTCTGGGGGAGGAGCAAGTTTAC 3849
QY |||||
QY 3907 TTCATGACTCTGAACCGTAACTGCATCATGAACCTGGTGA 3945
DB |||||
DB 3850 TTCATGACTCTGAACCGGTAACCTGCATCATGAACCTGGTGA 3888
```

RESULT 2

us-03-688-188b-11

```
Query Match 95.8%; Score 3785; DB 2; Length 4133;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 3824; Conservative 0; Mismatches 15; Indels 150; Gaps 3;

QY 110 CTTACGGACAGGTGTACAAAGGTCGGCATGTCAAGCGGGCAGCTGGCTGCATCAAGG 169
DB 2 CATTTGGGGAGGTGTATGAGGGTCGGCATGTCAAGCGGGCAGCTGGCTGCATCAAGG 61
QY 170 TCATGATGTCAACGAGGACGAGGAGGAGAGATCAAAACAGGAGATCAACATGCTGAAAA 229
DB 62 TCATGATGTCAACGAGGACGAGGAGGAGAGATCAAAACAGGAGATCAACATGCTGAAAA 121
QY 230 AGTACTCTCACACCGCAACATCGCCACCTACTACGAGCCCTTCATCAAGAGAGACCCCC 289
DB 122 AGTACTCTCACACCGCAACATCGCCACCTACTACGAGCCCTTCATCAAGAGAGACCCCC 181
QY 290 CGGGAACGATGACACGAGCTCTGGCTGGTGTGATGGAGTTCTGTGGTGTCTGTTTCAGTGA 349
DB 182 CGGGAACGATGACACGAGCTCTGGCTGGTGTGATGGAGTTCTGTGGTGTCTGTTTCAGTGA 241
QY 350 ACCTGGTAAAGAAACAAAAAGGCAACCGCCCTGAAGGAGGACTGTATCGCCTATATCTGCA 409
DB 242 ACCTGGTAAAGAAACAAAAAGGCAACCGCCCTGAAGGAGGACTGTATCGCCTATATCTGCA 301
QY 410 GGGAGATCCTCAGGGGTCTGGCCCATCTCCATGCCCAAGGTGATCCATCGAGACATCA 469
DB 302 GGGAGATCCTCAGGGGTCTGGCCCATCTCCATGCCCAAGGTGATCCATCGAGACATCA 361
QY 470 AGGGGAGATGTGCTGTGTGACAGAGATGCTGAGGTCAAGTCTGAGATTTTGGGGTGA 529
DB 362 AGGGGAGATGTGCTGTGTGACAGAGATGCTGAGGTCAAGTCTGAGATTTTGGGGTGA 421
QY 530 GTGCTCAGCTGGAACCGCACCGGTGGGCGAGACGGAACACTTTTCAATGGGACTCCCTACTGGA 589
DB 422 GTGCTCAGCTGGAACCGCACCGGTGGGCGAGACGGAACACTTTTCAATGGGACTCCCTACTGGA 481
QY 590 TGGCTCCAGAGGTCAATCGCTGTGTATGAGAACCCCTGTATGCCACCTATGATTTACAGAGTGT 649
DB |||||
```

Db 482 TGGCTCCAGAGTCAATCGCTGTGATGAGAACCCCTGATGCCACCTATGATTACAGGAGTG 541
Qy 650 ATATTGTGTTCTTAGGAATCACAGCCATCGAGATGGCAGAGGAGCCCCCTCTGTGTG 709
Db 542 ATATTGTGTTCTTAGGAATCACAGCCATCGAGATGGCAGAGGAGCCCCCTCTGTGTG 601
Qy 710 ACATGACCCCATCGAGGCCCTCTTCTCATTCCTCGGAAACCTTCGGCCAGGCTCAAGT 769
Db 602 ACATGACCCCATCGAGGCCCTCTTCTCATTCCTCGGAAACCTTCGGCCAGGCTCAAGT 661
Qy 770 CCAAGAAGTGTCTAAGAAGTTCAATGATTCATTGACACATGTCTCATCAAGACTTACC 829
Db 662 CCAAGAAGTGTCTAAGAAGTTCAATGATTCATTGACACATGTCTCATCAAGACTTACC 721
Qy 830 TGAGCGCCGCCACCCACGAGCAGCTACTGAAGTTTCCCTTCATCCGGGACCAAGCCACGG 889
Db 722 TGAGCGCCGCCACCCACGAGCAGCTACTGAAGTTTCCCTTCATCCGGGACCAAGCCACGG 781
Qy 890 AGCGGAGGTCCGATCCAGCTTAAGGACCAATTTGACCGATCCCGGAAGAGCGGGTG 949
Db 782 AGCGGAGGTCCGATCCAGCTTAAGGACCAATTTGACCGATCCCGGAAGAGCGGGTG 841
Qy 950 AGAAAGAGGAGACAGATATAGTACAGCGGAGGAGGAGGAGATGACAGCCATGGAG 1009
Db 842 AGAAAGAGGAGACAGATATAGTACAGCGGAGGAGGAGGAGATGACAGCCATGGAG 901
Qy 1010 AGAAAGGAGAGCCAAAGTCCATCATGAACGTGCTGGAGAGTCGACTTACCGCGGGAGT 1069
Db 902 AGAAAGGAGAGCCAAAGTCCATCATGAACGTGCTGGAGAGTCGACTTACCGCGGGAGT 961
Qy 1070 TTCTCGGCTCAGCAGGAAATAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGC 1129
Db 962 TTCTCGGCTCAGCAGGAAATAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGC 1021
Qy 1130 TGCAGCAGCAGCAGCAGAGACCCGAGGACACATCAAAACCTGTGTGACACAGCGGC 1189
Db 1022 TGCAGCAGCAGCAGCAGAGACCCGAGGACACATCAAAACCTGTGTGACACAGCGGC 1081
Qy 1190 AGCGGCGCATAGAGAGCAGAAAGAGAGCGCGCGCGTGGAGAGCAACAGCGCGCGG 1249
Db 1082 AGCGGCGCATAGAGAGCAGAAAGAGAGCGCGCGCGTGGAGAGCAACAGCGCGCGG 1141
Qy 1250 AGCGGAGCAGCGGAAGCTGCAGGAGAGGAGCAGCGCGGCTGCAGGACATGCAAG 1309
Db 1142 AGCGGAGCAGCGGAAGCTGCAGGAGAGGAGCAGCGCGGCTGCAGGACATGCAAG 1201
Qy 1310 CTCTCGGCGGAGGAGGAGCGCGGACAGCGGAGCTGACGA ----- 1352
Db 1202 CTCTCGGCGGAGGAGGAGCGCGGACAGCGGAGCGGAGCGGAGCAATATTGTCACG 1261
Qy 1353 ----- 1352
Db 1262 GGCTAGAGGAGGAGCAGCAGCTCGAGATCTTTCAGCAACAGCTGTCTCCAGGAAACAGG 1321
Qy 1353 -----GGAATACAGCGGAGCAGCTGGAGGAGCAGCGGAGCTCAGAACTCTCC 1402
Db 1322 CCCTGCTGTGTAATACAAAGCGGAAGCAGCTGGAGGAGCGCGGAGCTCAGAACTCTCC 1381
Qy 1403 AGAGGAGCTGCAGCAGGAGCAGCTTACCTCAAGTCCCTGAGCAGCAGCAACAGCAGC 1462
Db 1382 AGAGGAGCTGCAGCAGGAGCAGCTTACCTCAAGTCCCTGAGCAGCAGCAACAGCAGC 1441
Qy 1463 AGCAGCTTCAGAAAACAGCAGCAGCAGCTCTCTGCTGGGACAGGAAGCCCTGTACC 1522
Db 1442 AGCAGCTTCAGAAAACAGCAGCAGCAGCTCTCTGCTGGGACAGGAAGCCCTGTACC 1501
Qy 1523 ATTATGCTCGGGCATGAATCCCGCTGCAAAACAGCAGCTGGCGCGGAGAGGTAGAAGAGA 1582
Db 1502 ATTATGCTCGGGCATGAATCCCGCTGCAAAACAGCAGCTGGCGCGGAGAGGTAGAAGAGA 1561
Qy 1583 GAAACAGGATGAACAGCAGCAGAACTCTCTCTTGGCCAAAGCAGCAGCAGCAGCAGC 1642
Db 1562 GAAACAGGATGAACAGCAGCAGAACTCTCTCTTGGCCAAAGCAGCAGCAGCAGCAGC 1621

Qy 1643 GGCTTAGAGCCGCCCATCCCCAGGCTTCCCAAGGCCCCCAGAGACCCCTTTCCAGACTC 1702
Db 1622 GGCTTAGAGCCGCCCATCCCCAGGCTTCCCAAGGCCCCCAGAGACCCCTTTCCAGACTC 1681
Qy 1703 CTCCTATGACAGGCGGCTGAGCCCGCAGAGAGGAGCCGACAC ----- 1744
Db 1682 CTCCTATGACAGGCGGCTGAGCCCGCAGAGGAGCCGACACAGAGGCTGTGTGGCACC 1741
Qy 1745 -----AGTCCCTGCAGGACCCAGC 1762
Db 1742 GGGTCCCACTGAAGCCATATGCAGCACTGTATCCCGATCCCAGTCCCTGCAGGACCAGC 1801
Qy 1763 CCACCCGAAACTGGCTGCTTCCAGCTCCCATGACCCCGACCTTGCATCCCGGAC 1822
Db 1802 CCACCCGAAACTGGCTGCTTCCAGCTCCCATGACCCCGACCTTGCATCCCGGAC 1861
Qy 1823 CCATGCGCACGCGCAGTCCGAGGAGCTGTCAATCCGCCAGATTCAGAGCCCACTCTG 1882
Db 1862 CCATGCGCACGCGCAGTCCGAGGAGCTGTCAATCCGCCAGATTCAGAGCCCACTCTG 1921
Qy 1883 AAGGACCTGGCCCCCAGCCCGAATCCCCAGCTGGGTCCGCCAGATAAAGAGGCCCCAC 1942
Db 1922 AAGGACCTGGCCCCCAGCCCGAATCCCCAGCTGGGTCCGCCAGATAAAGAGGCCCCAC 1981
Qy 1943 CCAAGTGTCTCAGAGGACCTCATCTATCGGCACTGCCCCCTTAACACAGTGGGGCCGAG 2002
Db 1982 CCAAGTGTCTCAGAGGACCTCATCTATCGGCACTGCCCCCTTAACACAGTGGGGCCGAG 2041
Qy 2003 GGTCCCGCCAGCCAGGCACTCGTCCAGAGCTTCGACAGCACTCCGCTGGCAAACTC 2062
Db 2042 GGTCCCGCCAGCCAGGCACTCGTCCAGAGCTTCGAGCACTTCGAGCACTCCGCTGGCAAACTC 2101
Qy 2063 ATCTGCAAAAGCGCGGCAAGCGGGCACCCCAAGAGCTCCAGGCCCCCTTGTCTCAGCCCC 2122
Db 2102 ATCTGCAAAAGCGCGGCAAGCGGGCACCCCAAGAGCTCCAGGCCCCCTTGTCTCAGCCCC 2161
Qy 2123 CTGGCCCCCAAGCCCTCTAGTAAACCCGACCTCAGGAGGAGGAGCCCTGGCTGGGAAC 2182
Db 2162 CTGGCCCCCAAGCCCTCTAGTAAACCCGACCTCAGGAGGAGGAGCCCTGGCTGGGAAC 2221
Qy 2183 GCTCGGACAGCGCTCTTTCAGCCCTCTCAAGGCACTTCCCGGAGCTGGCTCACTGGAGC 2242
Db 2222 GCTCGGACAGCGCTCTTTCAGCCCTCTCAAGGCACTTCCCGGAGCTGGCTCACTGGAGC 2281
Qy 2243 GGAACCGCTGGAGCTCTTCCAAACTGACAGCTCCCTGTGCTCTCCCTTGGGAATA 2302
Db 2282 GGAACCGCTGGAGTCTCTCCAAACCGGACAGCTCCCTGTGCTCTCCCTTGGGAATA 2341
Qy 2303 AAGCCAAAGCCGACGACACCGCTCAGGCGCAGCGCCGCGAGACTTTGTGTGCTGA 2362
Db 2342 AAGCCAAAGCCGACGACACCGCTCAGGCGCAGCGCCGCGAGACTTTGTGTGCTGA 2401
Qy 2363 AAGAGCGGACTCTGGAAGGAGCCCTCGGCTTCCCAAGAGGCGCATGGAATCTCTGCTGT 2422
Db 2402 AAGAGCGGACTCTGGAAGGAGCCCTCGGCTTCCCAAGAGGCGCATGGAATCTCTGCTGT 2461
Qy 2423 CCAGCAGGAGGTGGAAGCAGTGAGGACGACAGAGGAGGAGGAGGAGGCGGCGCAGCAG 2482
Db 2462 CCAGCAGGAGGTGGAAGCAGTGAGGACGACAGAGGAGGAGGAGGAGGCGGCGCAGCAG 2521
Qy 2483 AGGGAGCAGAGATACCCCTGGGGCGCAGCGATGGGGATACAGACAGCGTTCAGACCA 2542
Db 2522 AGGGAGCAGAGATACCCCTGGGGCGC -----GCAATGGGGATACAGACAGCGTTCAGACCA 2578
Qy 2543 TGTGTGTCCAGCAGCTCGAGGAGATCACCGGAGCCAGCCCCCAATACGGGGCGGCAACA 2602
Db 2579 TGTGTGTCCAGCAGCTCGAGGAGATCACCGGAGCCAGCCCCCAATACGGGGCGGCAACA 2638
Qy 2603 TGTGTGTCCAGCAGCAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2662
Db 2639 TGTGTGTCCAGCAGCAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2698

QY	2663	ACACAAACCTGCTGCAGTGGTCCAGCCAGCCACTCACCCACCGAGAACACGAAGGCC	2722
Db	2699		
QY	2723	AAAGCCACCCCTCGAAGGATGGAGTGGTGACTACCACTCTGCTGGTGGTGAAGGCC	2782
Db	2759	AAAGCCACCCCTCGAAGGATGGAGTGGTGACTACCACTCTGCTGGTGGTGAAGGCC	2818
QY	2783	CTGGCAGAGCTCGTTCACGATGTTTGTGGATCTAGGGATCTTACCAAGCTCGAGGACAGT	2842
Db	2819	CTGGCAGAGCTCGTTCACGATGTTTGTGGATCTAGGGATCTTACCAAGCTCGAGGACAGT	2878
QY	2843	GGCAGACATCCCATCACAGCCCTAGTGGGTGGAGAGGGCACTCGGCTCGACCAAGCTGC	2902
Db	2879	GGGACAGCATCCCATCACAGCCCTAGTGGGTGGAGAGGGCACTCGGCTCGACCAAGCTGC	2938
QY	2903	AGTACGAGTGAAGGGTCTGTGGTCAACGTTGAATCCCAACCAACCCGGGCCACAC	2962
Db	2939	AGTACGAGTGAAGGGTCTGTGGTCAACGTTGAATCCCAACCAACCCGGGCCACAC	2998
QY	2963	GTGAGACCCCTGAGATCCGGAAGTACAGAGCGATTCACCTCCGAGATCCTCTGTGCAG	3022
Db	2999	GTGAGACCCCTGAGATCCGGAAGTACAGAGCGATTCACCTCCGAGATCCTCTGTGCAG	3058
QY	3023	CCCTTTGGGGGTCAACCTGCTGGTGGGCACGGAGAACGGGCTGATGTTGTGTCGACCGAA	3082
Db	3059	CCCTTTGGGGGTCAACCTGCTGGTGGGCACGGAGAACGGGCTGATGTTGTGTCGACCGAA	3118
QY	3083	GTGGGACGGGCAAGGTGTATGGACTCATTTGGGCGGCGACGTTTCCAGCAGATGGATGTC	3142
Db	3119	GTGGGACGGGCAAGGTGTATGGACTCATTTGGGCGGCGACGTTTCCAGCAGATGGATGTC	3178
QY	3143	TGAGGGGCTCAACCTGCTCATCAACCTCAGGGAAGGAAACAACTGCGGGTGTATT	3202
Db	3179	TGAGGGGCTCAACCTGCTCATCAACCTCAGGGAAGGAAACAACTGCGGGTGTATT	3238
QY	3203	ACCTGCTCGGCTCCGGAACAAGATTCTGCACATGACCAAGTGGAGAGAAAGCAGG	3262
Db	3239	ACTTGTCTGGCTCCGGAACAAGATTCTGCACATGACCAAGTGGAGAGAAAGCAGG	3298
QY	3263	GCTGGACCAACCTGGGGGACATGGAGGGCTGGGGCACTACCGTGTGTGAATAAGGAC	3322
Db	3299	GCTGGACCAACCTGGGGGACATGGAGGGCTGGGGCACTACCGTGTGTGAATAAGGAC	3358
QY	3323	GGATTAAAGTTCTGTGTCATCGCCCTCAAGACTCCGTGGAGGTGTATGCTGGGCCCCA	3382
Db	3359	GGATTAAAGTTCTGTGTCATCGCCCTCAAGACTCCGTGGAGGTGTATGCTGGGCCCCA	3418
QY	3383	AACCCCTACCAAAATTCATGGGCTTCAAGTCCCTTTCGCGACCTCCGCCACCGCCCTCTGC	3442
Db	3419	AACCCCTACCAAAATTCATGGGCTTCAAGTCCCTTTCGCGACCTCCGCCACCGCCCTCTGC	3478
QY	3443	TGCTCGACCTGACAGTAGAGAGGGGACGGCTCAAGTTCATCTATGCTCCAGTGTG	3502
Db	3479	TGCTCGACCTGACAGTAGAGAGGGGACGGCTCAAGTTCATCTATGCTCCAGTGTG	3538
QY	3503	GCTTCCATGCTGTGGATGTGACTCGGGGACAGCTATGACATCTACATCCCTGTGCACA	3562
Db	3539	GCTTCCATGCTGTGGATGTGACTCGGGGACAGCTATGACATCTACATCCCTGTGCACA	3598
QY	3563	TCAGAGCCAGATCAGCCCATGCCATCATCTTCTCCCAACACCGGACGGCATGGAGA	3622
Db	3599	TCAGAGCCAGATCAGCCCATGCCATCATCTTCTCCCAACACCGGACGGCATGGAGA	3658
QY	3623	TGCTGCTGTGTACGAGGACGAGGGTGTCTACGTCAACACGTAACGGGCGCATCATTAAGG	3682
Db	3659	TGCTGCTGTGTACGAGGACGAGGGTGTCTACGTCAACACGTAACGGGCGCATCATTAAGG	3718
QY	3683	ATGTGTGTGTGAGTGGGGAGATGCTACTTCTGTGGGCTACATCTGCTCCAAACGAGA	3742
Db	3719	ATGTGTGTGTGAGTGGGGAGATGCTACTTCTGTGGGCTACATCTGCTCCAAACGAGA	3778
QY	3743	TNATGGGCTGGGGTGAGAAAGCCATTGAGATCCGCTCTGTGGAGACGGGCCACCTCGACG	3802

Db	3779	TAATGGGCTGGGGTGAGAAAAGCCATTGAGATCCGCTCTGTGGACCGGGCCACCTCGACG	3838
QY	3803	GGGTCTTCATGSCACAAACGAGCTCAGAGGCTCAAGTTCTGTGTGAGCGGAATGACAAAG	3862
Db	3839	GGGTCTTCATGSCACAAACGAGCTCAGAGGCTCAAGTTCTGTGTGAGCGGAATGACAAAG	3898
QY	3863	TGTTTTTTTGGCTCAGTCCGCTCTGGGGGCGAGCCAAAGTTTACTTATGACTCTGAACC	3922
Db	3899	TGTTTTTTTGGCTCAGTCCGCTCTGGGGGCGAGCCAAAGTTTACTTATGACTCTGAACC	3958
QY	3923	GTAACCTGATCATGAACTGTGTGAAGGGC	3951
Db	3959	GTAACCTGATCATGAACTGTGTGAAGGGC	3987

Search completed: August 26, 2004, 15:39:57
Job time : 23 secs

THIS PAGE BLANK (USPTO)

Result No.	Query \$			DB	ID	Description
	Score	Match	Length			
1	6909	99.7	1332	1	M4K6 HUMAN	Q8n4c8 homo sapien
2	6484	93.3	1308	1	M4K6 MOUSE	Q9jm52 mus musculus
3	4336	62.6	1360	1	TN1K HUMAN	Q9uke5 homo sapien
4	4017.5	58.0	1239	1	M4K4 HUMAN	Q98h19 homo sapien
5	3951.5	57.0	1233	1	M4K4 MOUSE	P97820 mus musculus
6	3286	47.4	916	1	TN1K MOUSE	P83510 mus musculus
7	2646	38.2	1080	1	M115 CAEEL	Q23356 caenorhabdi
8	738	10.7	894	1	M4K3 HUMAN	Q81vth homo sapien
9	729	10.5	862	1	M4K3 RAT	Q944i2 rattus norv
10	715.5	10.3	487	1	STK4 HUMAN	Q13043 homo sapien
11	694.5	10.0	491	1	STK3 HUMAN	Q13188 homo sapien
12	693.5	10.0	669	1	HPO DROME	Q8t0s6 drosophila
13	691	10.0	847	1	M4K5 MOUSE	Q8bpm2 mus musculus
14	687	9.9	846	1	M4K5 HUMAN	Q9y4k4 homo sapien
15	681.5	9.8	426	1	STZ5 HUMAN	Q00506 homo sapien
16	675.5	9.7	426	1	STZ5 MOUSE	Q942w1 mus musculus
17	667	9.6	1080	1	NRK1 YEAST	P38692 saccharomyc
18	660.5	9.5	833	1	M4K1 HUMAN	Q92918 homo sapien
19	656.5	9.5	443	1	STZ4 HUMAN	Q9y6e0 homo sapien
20	653.5	9.4	819	1	M4K2 HUMAN	Q12851 h mitogen-a
21	644	9.3	968	1	STK4 HUMAN	Q94804 homo sapien
22	638	9.2	821	1	M4K2 MOUSE	Q61161 mus musculus
23	636.5	9.2	966	1	STK4 MOUSE	O55098 mus musculus
24	630.5	9.1	681	1	M4K3 MOUSE	Q99jp0 mus musculus
25	629	9.1	827	1	M4K1 MOUSE	P70218 mus musculus
26	620.5	9.0	1501	1	NIIN DROME	P10676 drosophila
27	552	8.0	471	1	S1D1 SCHPO	O14305 schizosacch
28	551	8.0	982	1	SULU CAEEL	P46549 caenorhabdi
29	548	7.9	490	1	SPS1 YEAST	P08458 saccharomyc
30	533.5	7.7	658	1	PAK1 SCHPO	P50527 schizosacch
31	507	7.3	971	1	CLM4 CANAL	O14427 candida alb
32	505	7.3	652	1	SPK1 SCHPO	O75011 schizosacch
33	504	7.3	556	1	SPAK MOUSE	Q921w9 mus musculus

Db 721 PLQRTSSGSSSSSTPSSQSGSQSGSQRTRVANSKSGSVLPHEPA-KV 779
QY 768 KPDHHR--SRGPPADFLV-----LKERTLEAPPPKPKAMDYSSSSSESVSS 813
Db 780 KPEESRDITRPPASPKAIBDLTALAKEUREIETNRMKPKVDTYSSSESS 839
QY 814 EDDEEGEGGPAAGS---RDTP-----GGRSGDSTDVSMTVMVHDVEEITGTQPPYGG-- 863
Db 840 EBEEDGESETHDGTAVASDIPRLIPTGAGSNEQYVGVGTHGLE--TSHADSPSGSI 897
QY 864 ----GTWVORTPEERNLHASNGVT----NLPDVVQPSHS-----PTENSKGQSPSKD- 912
Db 898 SREGTLMIRESGKRSKSHSDNGFAGHINLPDLVQSHSPAGTPTTEGLGRVSTHSORM 957
QY 913 GSGDYQSRGLVAPKAGKSSFTMFVDLGIYQPGS-----GDSIPITALVGEGTRLDQLQY 967
Db 958 DSGTEYGMG---SSYKASTFPFVDRPVYQTSPTDEDEDESSAAALFTSELLRQKAL 1014
QY 968 -DVRKGSVVNVNPTNRAHSETPEIRKYKRFENSEILCAALGVNLLVGTENGLMLLORS 1026
Db 1015 NEARKISVVNVNPTNIRPHSDTPEIRKYKRFENSEILCAALGVNLLVGTENGLMLLORS 1074
QY 1027 GQKVVGLIGRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPEYKQKQ 1086
Db 1075 GQKVVNLRNRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPEYKQKQ 1134
QY 1087 WTTVGDMECGHVRVYKRIKFLVIALKSSVEVYAWAPKPYHKMAFKSPADLPHRPLL 1146
Db 1135 WITVGDLECIYKVVYKRIKFLVIALKNAVEIYAWAPKPYHKMAFKSPADLQHKPLL 1194
QY 1147 VDLTVEEGORLKVYIGSSAGFAVDVDSGNSDIYIPVHIQIQTIPHAIIFLPNTDGMEM 1206
Db 1195 VDLTVEEGORLKVIFGSHGTGFHVDVDSGNSDIYIPSHIQIQTIPHAIVILPKTDGMEM 1254
QY 1207 LLYCDEBGVYNTYRIIKDVLQWGMPTSVAYICSNQIMGWGKAIEIRSVETGHLDG 1266
Db 1255 LVCYDEBGVYNTYGRITKDVVLQWGMPTSVAYIHSNQIMGWGKAIEIRSVETGHLDG 1314
QY 1267 VFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLRNRCIMXW 1312
Db 1315 VFMHKAQRLKFLCERNDKVFFASVRSGGSSQVFFMTLRNRCIMXW 1360

RESULT 4
M4K4 HUMAN
ID M4K4 HUMAN STANDARD; PRT; 1239 AA.
AC Q95819; O75172; Q9NS77;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 4 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase 4) (MEK kinase kinase 4) (MEKKK 4)
DE (HPK/GCK-like kinase HGK) (Nck interacting kinase).
GN NAKP4K4 OR HGK OR NIK OR KIAA0687.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 4 AND 5), FUNCTION, AND TISSUE SPECIFICITY.
RP SPECIFICITY.
RC TISSUE=Macrophage;
RX MEDLINE=99107863; PubMed=9890973;
RA Yao Z., Zhou G., Wang X.S., Brown A., Diener K., Gan H., Tan T.-H.;
RT "A novel human STE20-related protein kinase, HGK, that specifically activates the c-Jun N-terminal kinase signaling pathway.";
RL J. Biol. Chem. 274:2118-2125(1999).
RN [2]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Glialastoma;
RX MEDLINE=22499944; PubMed=12612079;
RA Wright J.H., Wang X., Manning G., LaMere B.J., Le P., Zhu S.,
RA Khatri D., Flanagan P.M., Buckley S.D., Whyte D.B., Howlett A.R.,

RA Bischoff J.R., Lipson K.E., Jallal B.;
RT "The STE20 kinase HGK is broadly expressed in human tumor cells and can modulate cellular transformation, invasion, and adhesion.";
RL Mol. Cell. Biol. 23:2068-2082(2003).
RN [3]
RP SEQUENCE OF 38-1239 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
RN [4]
RP SEQUENCE OF 262-1239 FROM N.A. (ISOFORM 3).
RC TISSUE=Testis;
RX Dueterhoeft A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 700-1239 FROM N.A.
RC TISSUE=Brain;
RA Saito T., Seki N., Hori T.;
RT "Isolation, expression profile and chromosome assignment of a novel serine/threonine kinase gene.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP INTERACTION WITH SPG3A.
RX MEDLINE=22275851; PubMed=12387898;
RA Luan Z., Zhang Y., Liu A., Man Y., Cheng L., Hu G.;
RT "A novel GTP-binding protein HGBP3 interacts with NIK/HGK.";
RL FEBS Lett. 530:233-238(2002).
CC -!- FUNCTION: Serine/threonine kinase that may play a role in the response to environmental stress and cytokines such as TNF-alpha. Appears to act upstream of the c-jun N-terminal pathway.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium.
CC -!- SUBUNIT: Interacts with the SH3 domain of the adapter proteins Nck (By similarity). Binds, via its CNH regulatory domain, to the N-terminal region of SPG3A.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1; Synonyms=Tumor-associated;
CC IsoId=O95819-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O95819-2; Sequence=VSP_007054, VSP_007057;
CC Name=3;
CC IsoId=O95819-3; Sequence=VSP_007056, VSP_007057;
CC Name=4; Synonyms=HGK-S;
CC IsoId=O95819-4; Sequence=VSP_007054, VSP_007055, VSP_007057,
CC Name=5; Synonyms=HGK-L;
CC IsoId=O95819-5; Sequence=VSP_007054, VSP_007055, VSP_007056, VSP_007057, VSP_007058;
CC -!- TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all tissue types examined. Isoform 5 appears to be more abundant in the brain, isoform 4 is predominant in the liver, skeletal muscle and placenta.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. STE20 subfamily.
CC -!- SIMILARITY: Contains 1 CNH domain.
CC -!- CAUTION: Ref. 4 sequence differs from that shown due to a frameshift in position 1151.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF096300; AAD16137.1; -.

Db 1178 EKAIEIRSIVETGHLGDGVFMHKAQRKFLCERNDKVFFASVRSGSSQVYFMTLQRTSLL 1237
Qy 1311 NW 1312
Db 1238 SW 1239

RESULT 5
M4K4 MOUSE
ID M4K4 MOUSE STANDARD; PRT; 1233 AA.
AC P97820;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase kinase 4 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase kinase 4) (MEK kinase kinase 4) (MEKKK 4)
DE (HPK/GCK-like kinase HGK) (Nck interacting kinase).
GN MAP4K4 OR NIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Brain, and Embryo;
RX MEDLINE=97280817; PubMed=9135144;
RA Su Y.-C., Han J., Xu S., Cobb M., Skolnik E.Y.;
RT "NIK is a new Ste20-related kinase that binds NCK and MEK1 and
RT activates the SAPK/JNK cascade via a conserved regulatory domain.";
RL EMBO J. 16:1279-1290(1997).
RN [2]
RP FUNCTION, AND INTERACTION WITH NCK.
RX MEDLINE=20136034; PubMed=10669731;
RA Becker E., Huynh-Do U., Holland S., Pawson T., Daniel T.O.,
RA Skolnik E.Y.;
RT "Nck-interacting Ste20 kinase couples Eph receptors to c-Jun N-
RT terminal kinase and integrin activation.";
RL Mol. Cell. Biol. 20:1537-1545(2000).
CC -!- FUNCTION: Serine/threonine kinase that may play a role in the
CC response to environmental stress and cytokines such as TNF-alpha.
CC Appears to act upstream of the c-jun N-terminal pathway.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium.
CC -!- SUBUNIT: Interacts with the SH3 domain of the adapter proteins
CC Nck. Binds, via its CNH regulatory domain, to the N-
CC terminal region of SPG3A (By similarity).
CC -!- TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all
CC tissue types examined. Highest levels observed in heart and
CC brain.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -!- SIMILARITY: Contains 1 CNH domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U88984; AAC53165.1; -.
DR PIR; T30989; T30989.
DR MGD; MGI:1349394; Map4k4.
DR GO; GO:0005524; P:ATP binding; IDA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IDA.
DR GO; GO:0006469; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0007243; P:protein kinase cascade; IDA.
DR GO; GO:0006950; P:response to stress; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.

Pfam; PF00780; CNH; 1.
Pfam; PF00669; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00036; CNH; 1.
SMART; SM00220; S_TK; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; 1.
PROSITE; PS00111; PROTEIN KINASE DOM; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 25 289 PROTEIN KINASE.
FT DOMAIN 920 1211 CNH.
FT NP_BIND 31 39 ATP (BY SIMILARITY).
FT BINDING 53 53 ATP (BY SIMILARITY).
FT ACT_SITE 152 152 BY SIMILARITY.
SQ SEQUENCE 1233 AA; 140601 MW; 869896D414459B19 CRC64;
Query Match 57.0%; Score 3951.5; DB 1; Length 1233;
Best Local Similarity 62.3%; Pred. No. 4.7e-113;
Matches 846; Conservative 112; Mismatches 228; Indels 173; Gaps 32;
Qy 1 MGDPAARSDDIDLSALRDPAGIFELVGVNGCTYGVYKGRHVKTQQLAAIKVMDVTE 60
Db 1 MANDSPAKSLVDIDLSLRDPAGIFELVGVNGCTYGVYKGRHVKT-VTAAIKVMDVTE 59
Qy 61 DEEEIKQEIINMLKYSHHRNIATYYGAFIKKSPGNDQDLWLWMEFCGAGSVTDLVKNT 120
Db 60 DEEEITLEINMLKYSHHRNIATYYGAFIKKSPGDDQDLWLWMEFCGAGSITDLVKNT 119
Qy 121 KGNALKSDCTAYICREILRGHLHAHKVHRDIKQNVLLTENAELVDFGVSAQLDR 180
Db 120 KGNLTKEDWIAYISREILRGHLHIIHVHRDIKQNVLLTENAELVDFGVSAQLDR 179
Qy 181 TVGRNTFIGTPYMAPEVIACDENPDATDYRSDIWSLGITATEMAEGAPPLCDMHPMR 240
Db 180 TVGRNTFIGTPYMAPEVIACDENPDATDYRSDLWSLGITATEMAEGAPPLCDMHPMR 239
Qy 241 ALFLIPRNPRLKSKKSKKFFIDITCLIKTLYSRPTEQLLKFFPIRQDQTERQVRI 300
Db 240 ALFLIPRNPRLKSKKSKKFFSIFGCLVKNTMQRPSTEQLLKHPFIRQDQNERQVRI 299
Qy 301 QLKDHIDRSKRKGEKETEYVSGSEEDDSHGE-EGEPSSIMNVGSESTLRREFLRQ 359
Db 300 QLKDHIDTRKKGKEDETEYVSGSEEEVEPEGEPEPSSIVNVFPESTLRDFLRQ 359
Qy 360 QENKSNSEALKQOQLOQOQOQDPEAHIKHLHORQRRIEIEKEERRRVEEQRRREOR 419
Db 360 QENKSEALRQQLQEQQLREGEYKQLLAERQRRIEQKQRRRLBEEQQRRREAR 419
Qy 420 KLQEKQ-----QRLEDQALRR-----EEERRQAEREQYKQKL-EQQRSE 463
Db 420 RQEREQRREQEKEKRLLELERRRKEEERARRAEEKRRVEREQEYIRRLQEEQRHLE 479
Qy 464 RLQRLQEQEHAYL--KSLQOQOQOQLOKQOQOQLLPQDR--KPLYHYGRGNPADPAPW 519
Db 480 ILQOQLLQEQAMLLHRRRPHAQOQPPPOQOQDRSKSFHAPKPHY----DPADR--- 532
Qy 520 AREVEERTMNKQON-SPLAKSKSGSTGPEPPIQASGPGPLSQTPPMQRPVPEQSGP 578
Db 533 AREVQWHLASLKNVSPVSRSHSFS-----DPS-----KPA 565
Qy 579 HKSLQDQFTRNLAAFPASHDPPDPAIPAPTATPSARGAVIRQNSDPTSEGGPSPNPAPW 638
Db 566 HHHLRSQ-----DPC-----PPSRSEGLSQSSDSKSE--VPEPTQAWS 602
Qy 639 RPD-NEAPPKVPQORTSSIALNTSGA-----GSRPAQAVRPRSRNSAQIYLQRAER 693
Db 603 RSDSEVPVPRVPTTSRSPVLSRDSPLQGGGQGNQAGQ-RNSTSSIEPRLLWERVEK 661
Qy 694 GTPKPPGPPAPQPPGPPNASSNPDLRRSDPGWERSDSVLPASHGHLPQAGSLERNRVGASS 753
Db 662 LVPRPGS--GSSSGSSNSGSGPG---SHPG-----SOSGGERFRVRS 701
Qy 754 KLDSSPVLSPGNKA---KPDHRS--RPGRPADPV-----LLKERTLDEAPRPKKAMDYS 804

```
Db 702 KSEGP--SPGESAAPKDDKKVFRSLKPGFVDLTLAKBLRAVEDVRPHKVTYIS 759
Qy 805 SSSEV-----ESSEDEEGEGCGPAGSRDTPGCR-----SGDTPSVSTVMVHDVEEITG 856
Db 760 SSSESGTTDEEEDVEGEGADDSGPDTRAASSPNLSNGETESVKTMTIIVHDDVESEP 819
Qy 857 TOPPGCGTVMVQRTPEBERNLHADSNGYTNLPDVQPSHPSTENSGQSPSPKDGSGD 916
Db 820 AMTPSKEGTLVQRT-----QSASST----- 840
Qy 917 YQSRGLVAPGKSSFTMFVDLGIYQGGGDSIPITALVG--GEGTRLDQLQYD--VRKGS 973
Db 841 -----LQHKSSSTFTPIDRLLQISPSGT-TVTSVVGSCDGLRPEAIRQDPTKGS 894
Qy 974 VVNVNPTNTRAHSEFTPEIRKYKRNSEILCAALGVNLLVGTENGLMLLDRSGGQVYG 1033
Db 895 VVNVNPTNTPQSDTPEIRKYKRNSEILCAALGVNLLVGTESGLMLLDRSGGQVYP 954
Qy 1034 LIGRRRFOQMDVLEGLNLLITISGQNKLRVYVYLSWLNKILHNDPEVEKKQGTWTVGDM 1093
Db 955 LISRRFOQMDVLEGLNLLVITSGKDKLRVYVYLSWLNKILHNDPEVEKKQGTWTVGDL 1014
Qy 1094 ECGCHRVVYKVERIKFLVIALKSSVEVYVAPKPYHKFMAFKSPADLRPLLDVLTVEE 1153
Db 1015 ECGCHRVVYKVERIKFLVIALKSSVEVYVAPKPYHKFMAFKSPADLRPLLDVLTVEE 1074
Qy 1154 GQRLKVIYGSAGFHAVDVSNGSYDIYIPVHIQSOITPHAIIFLPNTDGMELLCYDE 1213
Db 1075 GQRLKVIYGSAGFHAVDVSNGSYDIYIPVHIQSOITPHAIIFLPNTDGMELLCYDE 1134
Qy 1214 GYVNTYGRITKDVVLQNGEMPTSVAYICSNQIMNGEKAIEIRSVETGHLDGVPFMHKA 1273
Db 1135 GYVNTYGRITKDVVLQNGEMPTSVAYIRSNQTMNGEKAIEIRSVETGHLDGVPFMHKA 1194
Qy 1274 QRLKFLCERNDKVFASVRSGSSGVYFMTLNRCIMKW 1312
Db 1195 QRLKFLCERNDKVFASVRSGSSGVYFMTLGRTSLLSW 1233

RESULT 6
TNIK_MOUSE
ID TNIK_MOUSE STANDARD; PRT; 916 AA.
AC P83510;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Trf2 and NCK interacting kinase (EC 2.7.1.37) (Fragments).
GN TNIK OR KIAA0551.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaide I., Osato N., Saito R., Hume D.A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gusincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Peste G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shlmada K.,
```

```
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [2]
RP SEQUENCE OF 359-842 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
CC -1- FUNCTION: Stress-activated serine/threonine kinase that may play a
CC role in the response to environmental stress. Appears to act
CC upstream of the c-jun N-terminal pathway. May play a role in
CC cytoskeletal regulation (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- SUBUNIT: Interacts with TRAF2 and NCK (by similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P83510-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=P83510-2; Sequence=VSP_007351;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -1- SIMILARITY: Contains 1 CNH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK039113; BAC30241.1; -
CC EMBL; AK041777; BAC31061.1; -
CC EMBL; AK088459; BAC40365.1; -
CC EMBL; AK322306; BAC65588.1; -
CC MGD; MGI:2444917; C53000801Srik.
CC GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
CC GO; GO:0007243; P:protein kinase cascade; ISS.
CC GO; GO:0006445; P:regulation of translation; ISS.
CC GO; GO:0006950; P:response to stress; ISS.
CC InterPro; IPR001180; Citron.
CC InterPro; IPR00719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC ATP-binding; Serine/threonine-protein kinase; Transferase;
KW
```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; Z50029; CAA90344.1; --
 DR PIR; T27622; T27622.
 DR WormPep; ZC504.4; CE19311.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser Thr pkin AS.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 21 281
 FT NP_BIND 27 35
 FT BINDING 50 50
 FT ACT_SITE 151 151
 FT DOMAIN 757 1060
 SQ SEQUENCE 1080 AA; 120785 MW; 4DAB06070FE26F63 CRC64;

Query Match 38.2%; Score 2646; DB 1; Length 1080;
 Best Local Similarity 43.3%; Pred. No. 1e-73;
 Matches 590; Conservative 156; Mismatches 268; Indels 350; Gaps 35;
 10 LDDIDLSALRDPAGIFELVVGNGTYGVYGRHVKTGQLAAIKVMQVDEDEEIKOE 69
 6 LDEIDLNSLRDPAGIFELVVGNGTYGVYGRHVKTGQLAAIKVMQVDEDEEIKOE 65
 70 INMLKKYSHHRIATYGAFFKSP--PGNDQLWLVMEFCAGSVTDLVKTGNALKE 127
 66 INMLKKYSHHRIATYGAFFKSP--PGNDQLWLVMEFCAGSVTDLVKTGNALKE 125
 128 DCIAYICREILRGLAHKVIHRIKQNVLLTENAELVDFGSAQLDRVTGRNT 187
 126 EWAIYICREILRGLVHLQSKVIHRIKQNVLLTDSAEVLDVFGSAQLDKVTGRNT 185
 188 FIGTPYMAPEVIACDENPDATYDRSDIWSLIGTAIEMAEAPLCPDMHPRALFLIPR 247
 186 FIGTPYMAPEVIACDESPATYDSK-----ITALEMAEGHPPLCDMHPRALFLIPR 238
 248 NPPRLK-SKKWKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRQPTQVRQLKQHI 306
 239 NPPPKLKNKKWTKKFFETIETLVKYDHYRPTTGALLRHPFIKEQPHQETIRHSIKEHI 298
 307 DRSRKRGEKETETEYEGSEEDDSHGEEGEPSSIMN-----VPGESTLRRREFLRQ 359
 299 DNRN--RVKQDADYEYSGSEDEPSPNRRGSGWGIRODSESSMIPMDTLRKGFKQLQ 356
 360 QENKSNSEALKQO-QOLQOQQOQDPEAHIKHLHQRRRIEQQE-----ERRRVE 409
 357 ESSRGFAEPGAQQLRRLQPPAPAP-----PQYQQSRVVEPRRESSVKKLRAVSSRCAA 410
 410 EQQRERQRKLQEQEQQRRLD-----WQALRREERQAEQEYKQKQLEEQRS 462
 411 DGRHSPASRPRPVSHHRSQQSHAPAPHLADLANYEKKRRSRE----- 456
 463 ERLRQLQEQEHAYLSLQOQQOQQOQLQOQQOQLLPGDRKPLYHYRGWNPADKPWARE 522
 457 ERRERQAHAM----- 469
 523 VEERTRMNKQNSPLAKSGPGTGPEPPIPOASPGPGPLSGTQPPMWRP-----VEPOEG 577
 470 -----PIARVASVPAP-QQSRKWSSEPLITHVKPED- 500
 578 PHKSLOQDPTNLAAPFASHDPDPAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPPAW 637

501 -----LDVLASELSKMGCHH-----NGRSREE-----SMSPP-- 527
 638 VRPDNEAPPKVPQRTSSITATLANTSGAGSRPAQAVRARPRNSAWQIY--LQRAERGT 695
 528 -----PPAPPPREASISSITDIDVG-----ELDNGADAEMDDLKIMWNGE-GT 571
 696 PKPPGPPAQPPGPPNASSN--PDLRRSDPGWERSDSVLPASHGLPQAGSLERNVRVGASS 753
 572 LRGNKPL-PPTPTDGTENTLVSDVRNGNG-----NSGH-----GAYKGGKI----- 612
 754 KLDSSPVLSPGNKAKPDHRSRPGPADFVLLKERTLDEAPRPPKAMWDYSSSESEVESS 813
 613 -----PEIRPGIISLDDD-----DS 627
 814 EDDEEGEGGPAEGSRDTPGGRSDGTDVSTVMVHDVEBITGTQPPYGGGTVMVQRTPE 873
 628 DSDNEEG-----NEPL 638
 874 EERNLLHADNGYTNLPDVVQPSHSPTENSKGQSPKSGDYGQSRGLVKAPKSSP-- 931
 639 MFKP INASSRG--ALPDLLPKSPQLRRQINDQTRQMSDDRRADQPNQFQNSDRSRSTQH 696
 932 -----TMFVDLGIYQPGSGSDSIPTALVGEGT-----RLDQLQYDVRKGSV 975
 697 SFSNRDREKSFVG--YFGGAG-----AGGTVNRPRQDINQVQ-----V 736
 976 NVNPTN--TRAHSETPERIKYKRFNFSEILCAALWGVNLLVGTENGMLLDRSQGKVYG 1033
 737 NVTNSNGTPAENDAPAIRKYKKFSGEILCAALWGVNLLIGTDSGLMLLDRSQGKVYP 796
 1034 LIGRRRQOMDVLEGLNLLITISGRKNKLVYVLSWLRNKLHND-----PEVEKKQGW 1088
 797 LISRRRQDMTVLEGGQNILATISGRKIRIRVYLSWLRQKILRTEGAGSANTTEKRWG 856
 1089 TVGDMCGHVRVVKYERIKFLVIALKSSVEVYAWAPKPYHKWFAKSFADLPHRLVLD 1148
 857 NVGLQAGIHFVYERIKFLVVGLESEIETAWAPKPYHKWFAKSFADLPHRLVLD 916
 1149 LTVEEGRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSIQITPHAIIFLPTDGMEMLL 1208
 917 LTVEDNARLKVYSGTGFGFHAIDLDSAAVYDIYTPAQSGQTTTHCHVLPNSNGMQLL 976
 1209 CYDEGVVNTYGRRIKDVVLQMGEMPTSAVYICSNQIMGWGEKAIIRSVETGHLDGVF 1268
 977 CYDNEGVVNTYGRMTKXVVLQMGEMPTSAVYISTGOIMGNKAIIRSVETGHLDGVF 1036
 1269 MHKBAQRLKFLCERNDKVFPASVRSGGSSQVYFMTLNRNCIMNW 1312
 1037 MHKKAQKLKFLCERNDKVFPSSARGGSGCQIYFMTLNRKPLTNW 1080
 RESULT 8
 M4K3 HUMAN
 ID M4K3_HUMAN STANDARD; PRT; 894 AA.
 AC Q8IVH8; Q8IVH7; Q9UDMS; Q9Y6R5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
 DE (MAPK/ERK kinase kinase kinase 3) (MEK kinase kinase 3) (MEKKK 3)
 DE (Germinal center kinase related protein kinase) (GLK).
 GN MAP4K3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, TISSUE SPECIFICITY, AND
 RP MUTAGENESIS OF LYS-48.
 RC TISSUE=Macrophage, and Skeletal muscle;
 RX MEDLINE=97420743; PubMed=9275185;
 RA Diener K., Wang X.S., Chen C., Meyer C.F., Keesler G., Zukowski M.,

RA Tan T.-H., Yao Z.;
RT "Activation of the c-Jun N-terminal kinase pathway by a novel protein
RT kinase related to human germlinal center kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9687-9692(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RA Gorry M.C., Zhang Y., Marks J.J., Suppe B., Hart S., Cortelli J.,
RA Pallos D., Hart T.C.;
RT "Physical/genetic map of the 2p22-2p21 region on chromosome 2.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE OF 1-712 FROM N.A. (ISOFORM 1).
RA Edwards J., Wohlmann P., Hawkins M., Harkins R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: May play a role in the response to environmental stress.
CC Appears to act upstream of the c-jun N-terminal
CC pathway.
CC
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a
CC phosphorylated protein.
CC
CC -!- COFACTOR: Magnesium.
CC
CC -!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
CC MAPK3-mediated JNK activation (By similarity).
CC
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q8IVH8-1; Sequences=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=Q8IVH8-2; Sequences=VSP_007052;
CC Name=3;
CC IsoId=Q8IVH8-3; Sequences=VSP_007053;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues
CC examined, with high levels in heart, brain, placenta, skeletal
CC muscle, kidney and pancreas and lower levels in lung and
CC liver.
CC
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STP20 subfamily.
CC
CC -!- SIMILARITY: Contains 1 CNH domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AF000145; AAC15472.1; -.
DR EMBL; AF445413; AAN75849.1; -.
DR EMBL; AF445385; AAN75849.1; JOINED.
DR EMBL; AF445386; AAN75849.1; JOINED.
DR EMBL; AF445387; AAN75849.1; JOINED.
DR EMBL; AF445388; AAN75849.1; JOINED.
DR EMBL; AF445390; AAN75849.1; JOINED.
DR EMBL; AF445391; AAN75849.1; JOINED.
DR EMBL; AF445392; AAN75849.1; JOINED.
DR EMBL; AF445393; AAN75849.1; JOINED.
DR EMBL; AF445394; AAN75849.1; JOINED.
DR EMBL; AF445395; AAN75849.1; JOINED.
DR EMBL; AF445396; AAN75849.1; JOINED.
DR EMBL; AF445397; AAN75849.1; JOINED.
DR EMBL; AF445398; AAN75849.1; JOINED.
DR EMBL; AF445400; AAN75849.1; JOINED.
DR EMBL; AF445401; AAN75849.1; JOINED.
DR EMBL; AF445402; AAN75849.1; JOINED.
DR EMBL; AF445403; AAN75849.1; JOINED.
DR EMBL; AF445404; AAN75849.1; JOINED.
DR EMBL; AF445405; AAN75849.1; JOINED.
DR EMBL; AF445406; AAN75849.1; JOINED.
DR EMBL; AF445407; AAN75849.1; JOINED.
DR EMBL; AF445408; AAN75849.1; JOINED.
DR EMBL; AF445409; AAN75849.1; JOINED.
DR EMBL; AF445410; AAN75849.1; JOINED.
DR EMBL; AF445411; AAN75849.1; JOINED.
DR EMBL; AF445412; AAN75849.1; JOINED.
DR EMBL; AC007684; AAF19240.1; -.
DR HSSP; P24941; 1B38.
DR Genew; HGNC:6865; MAPK3.
DR MIM; 604921; -.
DR GO; GO:0005524; F:ATP binding; IDA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0007243; P:protein kinase cascade; IDA.
DR GO; GO:0006950; P:response to stress; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase;
KW Alternative splicing.
FT DOMAIN 16 273 PROTEIN KINASE.
FT NP_BIND 562 874 CNH.
FT NP_BIND 22 30 ATP (BY SIMILARITY).
FT BINDING 48 48 ATP.
FT ACT_SITE 136 136 BY SIMILARITY.
FT VARSPPLIC 1 12 MNPFGDLSRRNP -> MA (in isoform 2).
FT VARSPPLIC 352 372 Missing (in isoform 3).
FT FTID=VSP_007053.
FT FTID=VSP_007053.
FT MUTAGEN 48 48 K->S: LOSS OF KINASE ACTIVITY AND ABILITY
FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
SQ SEQUENCE 894 AA; 101315 MW; 6EB77BBB34E5B733 CRC64;
Query Match 10.7%; Score 738; DB 1; Length 894;
Best Local Similarity 29.1%; Pred. No. 3.2e-16;
Matches 200; Conservative 97; Mismatches 207; Indels 184; Gaps 22;
Qy 14 DLSALRDPAGIFELVGVNGTYGQVYKGRHVTGQLAAIKVMVDTEDEEIKQEIINML 73

DR EMBL; AF445409; AAN75849.1; JOINED.
DR EMBL; AF445410; AAN75849.1; JOINED.
DR EMBL; AF445411; AAN75849.1; JOINED.
DR EMBL; AF445412; AAN75849.1; JOINED.
DR EMBL; AF445413; AAN75850.1; -.
DR EMBL; AF445385; AAN75850.1; JOINED.
DR EMBL; AF445386; AAN75850.1; JOINED.
DR EMBL; AF445387; AAN75850.1; JOINED.
DR EMBL; AF445388; AAN75850.1; JOINED.
DR EMBL; AF445390; AAN75850.1; JOINED.
DR EMBL; AF445391; AAN75850.1; JOINED.
DR EMBL; AF445392; AAN75850.1; JOINED.
DR EMBL; AF445393; AAN75850.1; JOINED.
DR EMBL; AF445394; AAN75850.1; JOINED.
DR EMBL; AF445395; AAN75850.1; JOINED.
DR EMBL; AF445397; AAN75850.1; JOINED.
DR EMBL; AF445398; AAN75850.1; JOINED.
DR EMBL; AF445399; AAN75850.1; JOINED.
DR EMBL; AF445400; AAN75850.1; JOINED.
DR EMBL; AF445401; AAN75850.1; JOINED.
DR EMBL; AF445402; AAN75850.1; JOINED.
DR EMBL; AF445403; AAN75850.1; JOINED.
DR EMBL; AF445404; AAN75850.1; JOINED.
DR EMBL; AF445405; AAN75850.1; JOINED.
DR EMBL; AF445406; AAN75850.1; JOINED.
DR EMBL; AF445407; AAN75850.1; JOINED.
DR EMBL; AF445408; AAN75850.1; JOINED.
DR EMBL; AF445409; AAN75850.1; JOINED.
DR EMBL; AF445410; AAN75850.1; JOINED.
DR EMBL; AF445411; AAN75850.1; JOINED.
DR EMBL; AF445412; AAN75850.1; JOINED.
DR EMBL; AC007684; AAF19240.1; -.
DR HSSP; P24941; 1B38.
DR Genew; HGNC:6865; MAPK3.
DR MIM; 604921; -.
DR GO; GO:0005524; F:ATP binding; IDA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0007243; P:protein kinase cascade; IDA.
DR GO; GO:0006950; P:response to stress; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase;
KW Alternative splicing.
FT DOMAIN 16 273 PROTEIN KINASE.
FT NP_BIND 562 874 CNH.
FT NP_BIND 22 30 ATP (BY SIMILARITY).
FT BINDING 48 48 ATP.
FT ACT_SITE 136 136 BY SIMILARITY.
FT VARSPPLIC 1 12 MNPFGDLSRRNP -> MA (in isoform 2).
FT VARSPPLIC 352 372 Missing (in isoform 3).
FT FTID=VSP_007053.
FT FTID=VSP_007053.
FT MUTAGEN 48 48 K->S: LOSS OF KINASE ACTIVITY AND ABILITY
FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
SQ SEQUENCE 894 AA; 101315 MW; 6EB77BBB34E5B733 CRC64;
Query Match 10.7%; Score 738; DB 1; Length 894;
Best Local Similarity 29.1%; Pred. No. 3.2e-16;
Matches 200; Conservative 97; Mismatches 207; Indels 184; Gaps 22;
Qy 14 DLSALRDPAGIFELVGVNGTYGQVYKGRHVTGQLAAIKVMVDTEDEEIKQEIINML 73


```

Db 6 DLNR-RNQEDPELLQIRGSGYGVYKARNVNTGELAAIKVLEPGEDFALVQOEIIM 64
Qy 74 KYSHRNATYGAIFIKKSPGNDQDLWLMFECGAGSVTDLVKNTKGNALKEDEIAYI 133
Db 65 MKDCKHPNIVAFSGYLRR-----DKLWICMEFCGGSLQDIYHVT--GPLSELQIAYV 116
Qy 134 CREILRLAHLAHKVIHRDIKQONVLLTENAELVDFGVSAQLDRTVGRNTFICTPY 193
Db 117 SRETLQGLYLLHSGKMRDIIKGANILLTDNGHVKLADFGVSAQITATIAKRKSPIGTPY 176
Qy 194 WMAPEVIACDENPDATYDVSIDWLSGLITATEMAEAGAPLDCMHMPRALFLIPRN--PPP 251
Db 177 WMAPEVAVERK--GGYNQLCDLWAVGITALELAELQPPMFDLHPMALFLMTKSNFQPP 234
Qy 252 RLKSK-KWSKFIDFIDTCLIKTYLSRPPTEQLLKFPIRDPQTERQVRIQLKHIDRSR 310
Db 235 KLKDKWKSNSPHFHVMAKTKNPKRPTAEKLLQHPFVTOHLT-RSLAIELLDKVNPN- 292
Qy 311 KKRGEKETEYSSSEEDSHGEEGEPSSIMNVPGESTLRRFRLIQENKSNSEALK 370
Db 293 -----DHSYTHDFDD-----DPEPLVAVP----- 312
Qy 371 QQQQLQQQQQDPEAHIKHLHQRRRIEKEERRRVEEQORREORKEQKEQORRL 430
Db 313 -----HRIHSTGRNVREKTRSEITFGVKPDPPLRK--ETEPHHEL 352
Qy 431 EDMQA-LRREBE-----BQAREEQEYKR-----KQLEEQQRSELRQRL 469
Db 353 PSDGFLDSSEIYYTARNLQLEYGQHGQGYFLGANKSLKSVSEELHQHVAHL 412
Qy 470 QOEHALKSLQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO 529
Db 413 EDDEG-----DDDESKHSTLAKIPPLPPPKSIF-----IPQEMHS 450
Qy 530 NKQNSPLAKSPGSGTGPPIPOASPGPGPLSQTPMQRPVEPQEGPHKSLQDQPTRN 589
Db 451 TEDENQGTIKKCPMSGSGAKP-----SQVPP--RPPPLPPHPKV--ALNGG 494
Qy 590 LAAPFASHPDPPAIPAPTATPSARGAVIRQNSD-----PTSEGGPGSPN 633
Db 495 MSSFQLNGERD-----GSLCQQNEHRTGTLNRKEKDVKPKPSINGLPPTPK 541
Qy 634 -----PPAWVRPD 641
Db 542 VHMGAFCFSKVFNGCPLKIHCASSWINPD 569

```

RESULT 9

```

M4K3 RAT
ID M4K3 RAT STANDARD; PRT; 862 AA.
AC Q924I2,
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
DE (MAPK/ERK center kinase kinase 3) (MEK kinase kinase 3) (WEKK 3)
DE (Germinal center kinase related protein kinase) (GLK) (Fragment).
GN MAP4K3
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI TaxID=10116;
RN [1]_
RP SEQUENCE FROM N.A., AND INTERACTION WITH SH3GL2.
RX MEDLINE=21369947; PubMed=11384986;
RA Ramjaun A.N., Angers A., Legendre-Guillen V., Tong X.-K.,
RA McPherson P.S.;
RT "Endophilin regulates JNK activation through its interaction with the
RT germinal center kinase-like kinase.";
RL J. Biol. Chem. 276:28913-28919(2001);
CC -!- FUNCTION: May play a role in the response to environmental stress.
CC Appears to act upstream of the c-jun N-terminal pathway (By

```

```

similarity).
-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-!- COFACTOR: Magnesium (By similarity).
-!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
MAP4K3-mediated JNK activation.
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
STE20 subfamily.
-!- SIMILARITY: Contains 1 CNH domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL; AF312224; AAK53214.1; -.
HSSP; P24941; 1BUH
GO; GO:0005524; F:ATP binding; ISS.
GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
GO; GO:0004668; F:protein amino acid phosphorylation; ISS.
GO; GO:0007243; P:protein kinase cascade; ISS.
GO; GO:0006950; P:response to stress; ISS.
InterPro; IPR001180; Citron.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR008271; Ser_Thr_pkinase.
InterPro; IPR002290; Ser_Thr_pkinase.
Pfam; PF00780; CNH; 1.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00036; CNH; 1.
SMART; SM00220; S_TKC; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; 1.
PROSITE; PS00113; PROTEIN KINASE DOM; 1.
PROSITE; PS00108; PROTEIN KINASE ST; FALSE NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase.
FT NON_TER 1
FT DOMAIN 5 262 PROTEIN KINASE.
FT DOMAIN 530 842 CNH.
FT NP_BIND 11 19 ATP (BY SIMILARITY).
FT BINDING 34 34 ATP (BY SIMILARITY).
FT ACT_SITE 125 125 BY SIMILARITY.
SQ SEQUENCE 862 AA; 97390 MW; 58013AC3B0A3287F CRC64;

```

Query Match 10.5%; Score 729; DB 1; Length 862;

Best Local Similarity 30.4%; Pred. No. 5.8e-16;

Matches 200; Conservative 97; Mismatches 203; Indels 158; Gaps 24;

```

Qy 21 PAGIFELVEVNGTYGQYVKGRHVTKGQLAAIKMDVTEDEEEIKQEINMLKYSHR 80
Db 1 PQEDFELIQIGSGYGVYKARNVNTGELAAIKVLEPGEDFALVQOEIIMKCKHA 60
Qy 81 NIATYGAIFIKKSPGNDQDLWLMFECGAGSVTDLVKNTKGNALKEDEIAYICREILRG 140
Db 61 NIVAFSGYLRR-----DKLWICMEFCGGSLQDIYHVT--GPLSELQIAYVRETLQG 112
Qy 141 LAHLAHKVIHRDIKQONVLLTENAELVDFGVSAQLDRTVGRNTFICTPYWMAPEVI 200
Db 113 LYLLHSGKMRDIIKGANILLTDNGHVKLADFGVSAQITATIAKRKSPIGTPYWMAPEVA 172
Qy 201 ACDSNPATYDVSIDWLSGLITATEMAEAGAPLDCMHMPRALFLIPRN--PPPRLKSK-K 257
Db 173 AVERK--GGYNQLCDLWAVGITALELAELQPPMFDLHPMALFLMTKSNFQPPKLOK 230
Qy 258 WSKKFIDFIDTCLIKTYLSRPPTEQLLKFPIRDPQTERQVRIQLKHIDRSKRKGEKE 317
Db 231 WNSNFHFFVKNALTKNPKRPTAEKLLQHPFV-TQPLTRSLAIELLDKVNPN----- 281
Qy 318 ETEYEGSGSEEDSHGEEGEPSSIMNVPGEE--STLRREFRLQENKSNSEALKQOQOOL 375
Db 282 ----DHSYTHDFDD-----DPEPLVAVPHRIPSTSN-----VREKTRSEINFGQVKF 327

```

```
QY 376 OQQQQRDPPEAHXKHL-----HQRRRIEQQEERRRYVEQRRERQKQKEQ 426
Db 328 DPPLRKETEPH--HELDLQLEYQGHQSNYFLGKNKSLKVEEB-----LHQGH 376
QY 427 QRRLEDMQALRREREQRQERQEQYKQLEBQRSERLQRLQOEHAFLKSLQOQQQQ 486
Db 377 VAHLED-----DEGDD-----DSKGSTLKA----- 397
QY 487 QLQKQQOQLLPGDRKPLHYHGRMNPADKPAWAVEVERTRMNKQNSPLAKSPGSTG 546
Db 398 -----KVPPLPP-----KPK-SISIPDTHSSDSNOGTTIKRCPSSGS 435
QY 547 PEPPPIQASGPGPLSPQFPVQEQGPHKSLQDQPTNLAAFPASHDPDPAIPAP 606
Db 436 PAKP-SHVPFRPPP--RLPP-QKPAVLNG-----VSSFLNGERDGSVHQ 479
QY 607 TATPSARGAVIQNSD-----PTSEGGPSPN-----PPAWVRPD 641
Db 480 QSEQRGTNLRSRKEKDVPKPIGSLPPTPKVHMAGCFSKVFNCGPLKIHCAATSWINPD 537

RESULT 10
STK4_HUMAN STANDARD; PRT; 487 AA.
AC Q13043; Q15802; Q9NT24;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase
DE MST1) (MST-1) (Mammalian STE20-like protein kinase 1)
DE (Serine/threonine protein kinase Krs-2).
GN STK4 OR MST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95394929; PubMed=7665586;
RA Creasy C.L., Chernoff J.;
RT "Cloning and characterization of a human protein kinase with homology
RT to Ste20.";
RL J. Biol. Chem. 270:21695-21700(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413604; PubMed=8816758;
RA Taylor L.K., Wang H.C., Erikson R.L.;
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
RT 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
RN [3]
RP SEQUENCE OF 1-435 FROM N.A.
RA Laird G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
CC THE KINASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- PTM: Autophosphorylated on serine and threonine residues.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18297; AAA83254.1; -.
DR EMBL; U60207; AAB17262.1; -.

```

```
DR EMBL; AL109839; CAB89421.1; -.
DR HSSP; P24941; LHCL.
DR Genew; HGNC:11408; STK4.
DR MIM; 604965; -.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 30 281 PROTEIN_KINASE.
FT NP_BIND 36 44 ATP (BY SIMILARITY).
FT BINDING 59 59 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
FT DOMAIN 373 378 POLY-GLU.
FT CONFLICT 222 222 P -> R (IN REF. 1).
FT CONFLICT 312 312 V -> M (IN REF. 1).
SQ SEQUENCE 487 AA; 150758EBC5F77D5C CRC64;

Query Match 10.3%; Score 715.5; DB 1; Length 487;
Best Local Similarity 34.0%; Pred. No. 9.4e-16;
Matches 168; Conservative 98; Mismatches 155; Indels 73; Gaps 14;

QY 6 PARSLDDIDLSAL--RDPAGIFELVGVNGYGVYKGRVKTGQLAAIKMVDTEDEE 64
Db 10 PRRLKKLDEDSLTKQEEVDFVLEKLEGSGSYGVIKAIHKTQIIVAIKQVPVESLQ 69
QY 65 EIKOEINMLKKYSHRNIAITYGAFIKKSPGNDQDLWLMVMEFCGAGSVTDLVNKTGNA 124
Db 70 IIEISIMQCCDSPH--VVKYGSYFKNT-----DLWIMVEYCGAGSVSDIIR-LRNKT 120
QY 125 LKEDCIAIYCREILRGUHLAHLAHKVIHRDIKQNVLTENAEVKLVDFGVSQALDRTVGR 184
Db 121 LTEDEIATILQSTLKGLEYLHFMEKIHRIKAGNILLNTEGHAKLADFGVAGQLTDTMAK 180
QY 185 RNTFIGHTPYMAPEVIACDENPDATYDRSDIWSLIGITALEMARGAPLCLDMHPRMLFL 244
Db 181 RNTVIGTFPFWAPEVI-----QEGTNCVADINSIGITALEMARGKPPYADIHPMRAIFM 235
QY 245 IPRNPPRL--KSKKWSKKFIDFIDTCLIKTYLSRPTEQLLKFFIRDPQTERQVRIQLK 303
Db 236 IPTNPPPTFRKPELWSDNFTDFVKQCLVKSPQERATATQLLQHPFVR--SAKGVSI-LR 291
QY 304 DHIDRSKRKGEKEETEYEGSGSEEDDSDSHGESEGPSSIMNVPG----- 347
Db 292 DLINAEADVLRKQESQOREVDQDDSEENSEDEMDSGTMVRAVGDEMGTVRVASTMTDGA 351
QY 348 -----ESTLRREP--LRLLQKNKNSSEALKQOQLOOQ-----QQRDPPEAHXK 389
Db 352 NTMLIEHDTLPQGLGTWVINAEDEEGTWMKRRDRTMQPAKPSLFYFEQKEKENQNSF 411
QY 390 -----LLHQRRRIEKEKE-----ERRRYVEQRR-----ERQKQKEQKQRL 430
Db 412 GKSVPGLKNSDWWKIPQDGDYEFKSWTVDLQKRLALDPMMEQIEEIRQKYQSKRQ 471
QY 431 EDMAALRRERERRQ 444
Db 472 PILDAIEAKRRRQ 485

RESULT 11
STK3_HUMAN STANDARD; PRT; 491 AA.
ID STK3_HUMAN
AC Q13188; Q15445; Q15801;

```

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine protein kinase 3 (EC 2.7.1.37) (STE20-like kinase
 MST2) (MST-2) (Mammalian STE20-like protein kinase 2)
 DE (Serine/threonine protein kinase Krs-1).
 GN STK3 OR MST2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96144292; PubMed=8566796;
 RA Creasy C.L., Chernoff J.;
 RT "Cloning and characterization of a member of the MST subfamily of
 RT Ste20-like kinases.";
 RL Gene 167:303-306(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413604; PubMed=8816758;
 RA Taylor L.K., Wang H.C., Erikson R.L.;
 RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
 RT 2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
 RN [3]
 RP SEQUENCE OF 96-203 FROM N.A.
 RX MEDLINE=94100173; PubMed=8274451;
 RA Schultz S.J., Nigg E.A.;
 RT "Identification of 21 novel human protein kinases, including 3 members
 RT of a family related to the cell cycle regulator nima of Aspergillus
 RT nidulans.";
 RL Cell Growth Differ. 4:821-830(1993).
 CC -!- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT
 CC MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY
 CC SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,
 CC SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT
 CC HEART, LUNG AND BRAIN TISSUES.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC STE20 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U26424; AAC50386.1; -;
 DR EMBL: U60206; AAB17261.1; -;
 DR EMBL: 225422; CAAB0909.1; -;
 DR PIR: I38212; I38212.
 DR HSP: P24941; IHCL.
 DR Genew: HGNC:11406; STK3.
 DR MIM: 605030; -;
 DR GO: GO:0004674; P:protein serine/threonine kinase activity; TAS.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO: GO:0007165; P:signal transduction; TAS.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR008271; Ser thr pkin AS.
 DR InterPro: IPR002290; Ser thr kinase.
 DR Pfam: PF00069; pkinase; I.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding.
 KW DOMAIN 27 278 PROTEIN_KINASE.

FT NP_BIND 33 41 ATP (BY SIMILARITY).
 FT BINDING 56 56 ATP (BY SIMILARITY).
 FT ACT_SITE 146 146 BY SIMILARITY.
 FT DOMAIN 308 314 POLY-GLU.
 FT DOMAIN 370 375 POLY-GLU.
 FT CONFLICT 96 98 WIV -> YLY (IN REF. 3).
 FT CONFLICT 121 121 D -> Y (IN REF. 3).
 FT CONFLICT 203 203 D -> G (IN REF. 3).
 FT CONFLICT 303 303 D -> E (IN REF. 2).
 FT CONFLICT 332 334 GEC -> ESV (IN REF. 2).
 SQ SEQUENCE 491 AA; 56261 MW; 9CA3B064F3C14A9 CRC64;
 Query Match 10.0%; Score 694.5; DB 1; Length 491;
 Best Local Similarity 33.6%; Pred. No. 4.1e-15;
 Matches 173; Conservative 98; Mismatches 183; Indels 61; Gaps 16;
 QY 4 PAPARSLDDLSAL-RDPAGIFELVEVNGTYGYQVYKGRHVKTGQLAAIKVMVDTEDE 62
 DB 5 PAKSKLKLSEDSLTQKPEEVPDLEKLGSGSYGVFAIKHESGQVVAIKQPVESDL 64
 QY 63 EEEIKQEIIMLKYSHRNIATYTGAFIKSPGNDQDLWLVMEFCGAGSSTDVKNYTKG 122
 DB 65 QEIIK-BISIMQ-CDSPPYVVKYGSYFKN- - - - -DLWIVMEYCGAGSVSDIIR-LRN 115
 QY 123 NALKEDCIATICREILRLGLAHKLVHRIKIGQNVLLTENAELVGLVDFGVSAQLDRTV 182
 DB 116 KTLIEDIATILKSTLKGLEYLHPMKIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTM 175
 QY 183 GRRNTFTGTPYMAPEVIACDENPDATYDRSDIWSLGTITAIENAEAGAPPLCDMHPREAL 242
 DB 176 AKRNVITGTFMAPEVI - - - - -QEIGYNCVADISLIGTISIEAEGKPPYADIHMPRAI 230
 QY 243 FLIPRNPPLRL-KSKWKSXKPIDFTCLIKTYLSRPTTEOLLKPPFI - - - - -RDQ 292
 DB 231 FMIPNPPPTFRKPELWSDPTDFVKKCLVKNPQRATATQLLQHPFIKNAKPVSIIRD 290
 QY 293 PTERQVRIQLKDHDRSKRKEETEYEVSGSEEDDSH - - - - -GEES-EPSSIM 343
 DB 291 ITE-AMEIKAKRHDEQRELEEEESD - - - - -EDELDSHTVMVKTSGEGCTWRATSTM 343
 QY 344 NVPCESTRERFLRLOENKS - - - - -NSEALKQOQOQOQOQOQOQOQOQOQOQOQ 400
 DB 344 SEGATWIEHNSTWLESDLTGVTWVNSEDEEDGTMRKATSPQVQRFSDY - - - - -PDK 399
 QY 401 QKEERRVBEQRRERQRKLEQK - - - - -EQORLEDMQALRRERERERERERERER 452
 DB 400 QDFKNKSHENCNQNHPFPMSKNVFPDNNKVPQDGDGDFLKNLSLELQMLKALDPM 459
 QY 453 RKQLEERQSERLQRLQEHAYLKSLOQOQOQOQOQ 487
 DB 460 EREIEELRQRYTAKRQ - - - - -PILDAMDAKKRQO 489
 RESULT 12
 HPO_DROME
 ID_HPO_DROME STANDARD; PRT; 669 AA.
 AC Q8T0S6; Q9V8W4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serine/threonine protein kinase hippo (EC 2.7.1.37).
 GN HPO OR CG11228.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Metopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.; FUNCTION, TISSUE SPECIFICITY, AND INTERACTION WITH
 RP SAV AND WTS.
 RX MEDLINE=22841274; PubMed=12941274;
 RA Harvey K.F., Pfeleger C.M., Hariharan I.K.;
 RT "The Drosophila Mst ortholog, hippo, restricts growth and cell

RT proliferation and promotes apoptosis.";
RL Cell 114:457-467(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannk C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupezki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP REVISIONS.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley; TISSUE=Head;
RA MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=21898525; PubMed=11900973;
RA Brody T., Stivers C., Nagle J., Odenwald W.F.;
RT "Identification of novel Drosophila neural precursor genes using a
RT differential embryonic head cDNA screen."
RL Mech. Dev. 113:41-59(2002).
RN [6]
RP INDUCTION.
RX MEDLINE=22085919; PubMed=12091301;
RA Egger B., Leemans R., Loop T., Kammermeier L., Fan Y., Radimerski T.,
RA Strahm M.C., Certa U., Reichert H.;
RT "Glucogenesis in Drosophila: genome-wide analysis of downstream genes
RT of glial cells missing in the embryonic nervous system."
RL Development 129:3295-3309(2002).
RN [7]
RP FUNCTION, PHOSPHORYLATION OF SAV AND WTS, MUTANT HPO42-47, AND
RP MUTAGENESIS OF LYS-71.
RX MEDLINE=22822128; PubMed=12941273;
RA Wu S., Huang J., Dong J., Pan D.;
RT "hippo encodes a Ste-20 family protein kinase that restricts cell
RT proliferation and promotes apoptosis in conjunction with Salvador and
RT warts".
RL Cell 114:445-456(2003).
CC -1- FUNCTION: Serine/threonine kinase required to restrict cell
CC proliferation and promote apoptosis. Phosphorylates Sav and Wts
CC proteins. The interaction with Sav promotes Wts phosphorylation.
CC Probably part of a same tumor suppression pathway with Sav and
CC Wts. Regulates the level of Th/DAPI apoptosis inhibitor.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: Expressed in CNS during embryogenesis. In
CC third instar larvae, it is expressed throughout all imaginal
CC disks.
CC -1- INDUCTION: Transcriptionally regulated by Gcm (Glial cells
CC missing).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib.ch).
CC EMBL; AE003795; AAF57543.2; -;
DR EMBL; AY069088; AAL39233.1; -;
DR HSSP; P24941; LHCL
DR Flybase; FBgn034453; hpo.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase, 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Apoptosis; ATP-binding.
FT DOMAIN 42 293
FT NP BIND 48 56
FT BINDING 71 71
FT ACT SITE 161 161
FT DOMAIN 490 503
FT DOMAIN 538 545
FT DOMAIN 490 599
FT MUTAGEN 71 71
FT MUTAGEN 166 171
FT CONFLICT 503 503
SQ SEQUENCE 669 AA; 75109 MW; B3DDDFAP5F4E4EB3 CRC64;
Query Match 10.0%; Score 693.5; DB 1; Length 669;
Best Local Similarity 31.3%; Pred. No. 5.6e-15;
Matches 193; Conservative 104; Mismatches 203; Indels 117; Gaps 22;
Qy 16 SALRDPAGIELVEVNGVGTGYKGRHVKTGLAALKVMDVTDEDEEEIKQELNMLKK 75
Db 33 SLQPPKPVDMYKLGEGSYGVKAVHKESSVIAIKLVFVESDLHLIILK-EISIMQQ 91


```
DR InterPro: IPR008271; Ser thr_pkin_AS.  
DR InterPro: IPR002290; Ser thr_pkinase.  
DR InterPro: IPR001245; Tyr_pkinase.  
DR Pfam: PF00780; CNH; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Prot_kinase; 1.  
DR SMART: SM00036; CNH; 1.  
DR SMART: SM00220; S_TK; 1.  
DR SMART: SM00219; TyRK; 1.  
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.  
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.  
DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.  
KW ATP-binding; Transferase; Serine/threonine-protein kinase;  
KW Alternative splicing.  
FT DOMAIN 20 277  
FT NP_BIND 26 34  
FT BINDING 49 49  
FT ACT_SITE 140 140  
FT DOMAIN 501 827  
FT VARSPPLIC 294 312  
FT  
FT CONFLICT 79 79  
FT CONFLICT 103 103  
FT CONFLICT 118 118  
FT CONFLICT 441 441  
FT CONFLICT 727 727  
SQ SEQUENCE 847 AA; 95044 MW; 377F0BF6CEECFA3 CRC64;  
  
Query Match 10.0%; Score 691; DB 1; Length 847;  
Best Local Similarity 36.0%; Pred. No. 8.1e-15;  
Matches 169; Conservative 85; Mismatches 149; Indels 66; Gaps 15;  
  
QY 5 APARSLDDLSALRDPAGIFELVEGVNGTYGVYKGRHVNTGQLAAIKWMDVTE-DEE 63  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
3 APLRPAADI---LRNPQHDYELVORVGSYGVYKARNVHTGELAAVKKLEPGDDF 59  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
64 EBIQEIINLKKYSHRNATYYGAFIKKSPGNDQLWLWMEFCGAGSVTDLVNKTGN 123  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
60 SLIQEIEFVKE-CKHCNIVAYFGSYLSR-----EKLWICMEYCGGSLQDIYHVT--G 110  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
124 ALKEDCIAYICEILRLGLAHKVIHRDTKQNVLTATENAVKLVDPGSAQLDRTVG 183  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
111 PLUSEQIAVVCRTIQLGLAYLTKGMHRDITKGANILLTDHGDVKLADFGVAAKITATIA 170  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
184 RNTFTIGTPYMAPEVIAEDENPDATYDYSIDSLGTLATEMARGAPLCLDMHPMRALF 243  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
171 KEKSFIGTPYMAPEVAEVEKN--GGYNQLCDIWAAGITALELQELPMPDLHPMRALF 228  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
244 LIPRN--PPRLKSK-KWKKFIDFTCLIKTYLSRPPTQLKFPFIRDPQTERQVRI 300  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
229 LMSKSNFQPKLKDRTKWSSTFHNFKVKTALTKNPKRPTAERLLTHTFVGQPLSRALAV 288  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
301 QLKDHIDRSRKRGEKEETEYYSSEEDDSHGEEGPSSIMVPGESTLRREFLRLOQ 360  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
289 ELLDKV-----SNPDHAPYSGDEDD-----LEPHAIIRHTIRSTNRN-----SR 329  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
361 ENKNSSEALKKOOQLQOQQOQDPEAHIK-----HLHL-----QRRRIE 399  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
330 AERTASEINFQKLEPPLRKEATEADEMGLSSEPNFILWNPFDVGANTGRTSKRAIP 389  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
400 EQKEERRRVEEQRRERSRQKLEKEQQ-----RRLEDMQALRRREERRQA 445  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
390 PPLPKPRVNTY-----PEDSLPDEKSTIKRCPDLERAPQVLRQOS 433  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
  
RESULT 14  
M4K5 HUMAN  
ID M4K5 HUMAN STANDARD; PRT; 846 AA.  
AC Q9Y4K4; Q8IYF6;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 5 (EC 2.7.1.37)
```

```
DE (MAPK/ERK kinase kinase kinase 5) (MEK kinase kinase 5) (MEKKK 5)  
GN (Kinase homologous to SPK1/STE20) (KHS).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE  
RP SPECIFICITY, AND MUTAGENESIS OF LYS-49.  
RC TISSUE=T-cell;  
RX MEDLINE=97190217; PubMed=9038372;  
RT "A novel human SPK1/STE20 homologue, KHS, activates jun N-terminal  
RT kinase.";  
EL Oncogene 14:653-659 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bonak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield A.S., Krzywinski M.I., Skalska U., Smalusz D.B.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP INTERACTION WITH CRK AND CRKL.  
RX MEDLINE=99002481; PubMed=9788432;  
RA Oehrl W., Kardinal C., Ruf S., Adermann K., Groffen J., Feng G.-S.,  
RA Blenis J., Tan T.-H., Feller S.M.;  
RT "The germinal center kinase (GCK)-related protein kinases HPK1 and KHS  
RT are candidates for highly selective signal transducers of Crk family  
RT adapter proteins.";  
RL Oncogene 17:1893-1901 (1998).  
CC -|- FUNCTION: May play a role in the response to environmental stress.  
CC Appears to act upstream of the c-jun N-terminal pathway.  
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -|- COPACTOR: Magnesium.  
CC -|- SUBUNIT: Interacts with both SH3 domains of the adapter proteins  
CC CRK and CRKL.  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -|- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues  
CC examined, with high levels in the ovary, testis and prostate.  
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC STE20 subfamily.  
CC -|- SIMILARITY: Contains 1 CNH domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL; U77129; BAB48435.1; -;  
CC EMBL; BC036013; BAB36013.1; -;  
CC HSPF; P24941; 1B38.
```


DR Genew; HGNC:6867; MAP4K5.
DR MIM; 604923; -.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005524; F:ATP binding; IDA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IDA.
DR GO; GO:0007257; P:activation of JUNK; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0007243; P:protein kinase cascade; IDA.
DR GO; GO:0006950; P:response to stress; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR008271; Ser Thr_kin_AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase.
FT DOMAIN 20 277 PROTEIN_KINASE.
FT NP_BIND 26 34 ATP (BY SIMILARITY).
FT BINDING 46 49 ATP.
FT ACT_SITE 140 140 BY SIMILARITY.
FT DOMAIN 512 826 CNH.
FT MUTAGEN 49 49 K->R: LOSS OF KINASE ACTIVITY AND ABILITY TO ACTIVATE JNK FAMILY.
FT CONFLICT 473 473 N -> K (IN REF. 2). CRC64;
FT SEQUENCE 846 AA; 95039 MW; B17F95196AA14E08 CRC64;

Query Match 9.9%; Score 687; DB 1; Length 846;
Best Local Similarity 29.2%; Pred. No. 1.1e-14;
Matches 190; Conservative 100; Mismatches 217; Indels 144; Gaps 21;

QY 5 APARSLDIDLSALRDPAGIFELVEVNGNYGVYKGRHVKYTGQLAIAKWDVTE-DEE 63
DB 3 APLRPAADI---LRNPQODYELVQVSGTYGVYKARNVHTGELAIAVKIKLEPGDDF 59

QY 64 EEIKQENMLKYSHRNIATYGAIFIKKSPGNDQLWLVMEFCGAGSVTLVKNTKGN 123
DB 60 SLIQEIEFWKE-CHCNVAYFGSLR-----EKLWICMEYCGGSLQDIYHVT--G 110

QY 124 ALKEDCIAYICREILRLGLAHLAHKVIHRIKDGQNVLTENAELVKLVDFGVSQAQDRVTG 183
DB 111 PLSELQIAVCRETLQGLAYLHTKGMHRIKGANILITDHDVKLADFGVAAKITATIA 170

QY 184 RNTPTIGTPYMAPEVIAACDENPDATYDYSIDSLGITAEMAGAPLDCMDHPMRALF 243
DB 171 KRKSPIGTPYMAPEVAEAKN--GGYNQLCDIWAAGVITAEIELGELQPMFEDLHPMRALF 228

QY 244 LIPRN--PPRLKSK-KSKKFEIDFIDCLIKTYLSRPTQLLKFPPIRQOPTQVRI 300
DB 229 LMSKSNFQPKLQKTKSSFTFNFKIALTKNPKKRTAERLLTHTVAPGLSRALAV 288

QY 301 OLKHIDHRSKRKRGEKEETEYSGSEEDSHGEGEPPSSIMNVPGESTLRREFLRQQ 360
DB 289 ELLDKVN-----NPDNHAHYTEADDD----- 310

QY 361 ENKNSSEALKKQQQQLQOQQQRDPENAIKHLHLQRORRIEQQEERRRVEQORREQRK 420
DB 311 -----PEPFAIRHTI-----RSTNRARAERTASE 336

QY 421 LOEKEQQRLEDQMALREBEERQAEQYKRLERQSRERLQRLQOEHAVLKSLQ 480
DB 337 INFQ---KLOPEPLRLKETEARD-----EMGLSSDPNFMQLQWN 371

QY 481 OQOQQQQQLQKQQQLQKQDRKPLHYHGRGMNPKAPAWAEVERETRMNKQONSPLAKS 540
DB 372 FVVDGANTGKSTKRAIPPLPP-----KPRISSVPED--NFPDEKASTIKH 417

QY 541 KPGSTGPEPPIFQASFGPP--GPLSQTPPMQRPVEPQEGPHKSLQDQPTNLAA--FPASH 597

Db 418 CPDSESAPQILRRQSSFCGPVAET-----SSINGDGIKLSMSENTGSAQAQLPRKN 473
QY 598 D----PDPAIPAPTATESA-RGAVIRONSDPTSEGPSPNPPAWVRPDNE 643
Db 474 DKRDFPKPAINGLPPTPKVLWGACFSKVFDCG---PLKINCATSWIHPTDK 521

RESULT 15
ST25 HUMAN STANDARD; PRT; 426 AA.
AC O00506; Q15522;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase 25 (EC 2.7.1.37) (Sterile 20/oxidant stress-response kinase 1) (Ste20/oxidant stress response kinase-1) (SOK-1) (Ste20-like kinase).
DS STK25 OR SOK1 OR YSK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97042345; PubMed=8887545;
RA Pombo C.M., Bonventre J.V., Molnar A., Kyriakis J., Force T.;
RA "Activation of a human Ste20-like kinase by oxidant stress defines a novel stress response pathway.";
RL EMBO J. 15:4537-4546(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97304522; PubMed=9160885;
RA Osada S.-I., Izawa M., Saito R., Mizuno K., Suzuki A., Hirai S.-I., Ohno S.;
RA "YSK1, a novel mammalian protein kinase structurally related to Ste20 and SPK1, but is not involved in the known MAPK pathways.";
RL Oncogene 14:2047-2057(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smallos D.B., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ENZYME REGULATION: Activated by phosphorylation, probably autophosphorylation. The C-terminal non-catalytic region inhibits the kinase activity.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Highest levels are found in testis, large intestine, brain and stomach followed by heart and lung.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.

```

CC STE20 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99325; CAA67700.1; -
DR EMBL; D63780; BAA20420.1; -
DR EMBL; BC007852; RAH07852.1; -
DR PIR; S71886; S71886.
DR HSP; P12931; 1FMK.
DR Genew; HGNC:11404; STK25.
DR MIM; 602255; -
DR GO; GO:0004672; F:protein kinase activity; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 20 270 PROTEIN KINASE.
FT NP_BIND 26 34 ATP (BY SIMILARITY).
FT BINDING 49 49 ATP (BY SIMILARITY).
FT ACT_SITE 140 140 BY SIMILARITY.
FT CONFLICT 347 348 EP -> DA (IN REF. 1).
SQ SEQUENCE 426 AA; 48111 MW; 183CE5700FCEA716 CRC64;

Query Match 9.8%; Score 681.5; DB 1; Length 426;
Best Local Similarity 38.3%; Pred. No. 8.9e-15;
Matches 169; Conservative 67; Mismatches 124; Indels 81; Gaps 15;

QY 20 DPAGIFELVEVNGTYGVYKGRVHTKGLAAIKVMDV--TEDEEERIKOEINMLKKYS 77
DB 15 DEEELFTKLDRIKGSFGSEVYKGINHTKVEVAIKIIDEAEDEIEDIQEITVLSQ-C 73

QY 78 HHRNIATYGFAPFKKSPGNDQDLVLMVFCGAGSVTDLVKNTKGNALKEDCIAYICREI 137
DB 74 DSPYITRYFGSYLKST-----KLWIMEYLGGSALDILL--XGPLEEYVIATILREI 124

QY 138 LGLAHLAHAKVIRHDIKQNVLLTENAIEVKLVDFGSAQLDRTVGRRTFTGTPYMAP 197
DB 125 LKGLDYLHSEKIRHDIKAANVLLSEQGDVKLADFGVAGQLTDTQIKRNTFTVGFPMWAP 184

QY 198 EVIACDENPDATYDYSRDIWSLIGITAIEMAGCAPPLCDMPMRALFLIPRNPPLKSKX 257
DB 185 EVI-----KQSYDFADLWSLIGITAIELAKGPPNSDLHPMRVFLIPKNSPPTLEGQH 239

QY 258 WSKFIDFTDCLIKTYLSRPTEQLKFPFIROQTERQVRIQLKDHIDRSKRKRGEKE 317
DB 240 -SKPPKEFVEACLNKPRFPRTAKELLKHKFTITRYTKTSFTTEL---IDRYKR---WKS 292

QY 318 ETEYEYSGEED-DSHGEEGEPSSIMNVPGESTLRREFRLQENKNSALKKQQOOLQ 376
DB 293 EGHGESESDSDIDGEAEDGEGQPIWTFP--PTIR----- 326

QY 377 QQQQRDPEAHIGLHQRRIEEOKEERRRVEEQRREREQ-----RKLQEKQ 426
DB 327 -----PSFHSK--LHKGTALHSSQK-----AEVVKRQPRSQCLSTLVRPVFGLKKEHK 374

QY 427 Q-----RRLEDQALRR 439

```

Db 375 QSGGSGVGALELENAFSLAE 395

Search completed: August 28, 2004, 00:57:36
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:54:25 ; Search time 35 Seconds
(without alignments)
1935.236 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAAPSLDIDLSALRD.....SGGSSQVYFMTLNRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6659	96.1	1326	4	US-09-688-188B-15
2	6659	96.1	1326	4	US-09-291-417D-15
3	4370.5	63.1	1324	4	US-09-645-456A-13
4	4370.5	63.1	1324	4	US-09-425-324A-13
5	4370.5	63.1	1324	4	US-09-645-791-13
6	4360.5	62.9	1332	4	US-09-645-456A-9
7	4360.5	62.9	1332	4	US-09-425-324A-9
8	4360.5	62.9	1332	4	US-09-645-791-9
9	4346	62.7	1353	4	US-09-645-456A-11
10	4346	62.7	1353	4	US-09-425-324A-11
11	4346	62.7	1353	4	US-09-645-791-11
12	4336	62.6	1360	3	US-09-393-569-2
13	4336	62.6	1360	4	US-09-579-664B-14
14	4336	62.6	1360	4	US-09-645-456A-34
15	4336	62.6	1360	4	US-09-425-324A-34
16	4336	62.6	1360	4	US-09-645-791-34
17	4274	61.7	1269	4	US-09-645-456A-15
18	4274	61.7	1269	4	US-09-425-324A-15
19	4274	61.7	1269	4	US-09-645-791-15
20	4264	61.5	1277	4	US-09-645-456A-12
21	4264	61.5	1277	4	US-09-425-324A-12
22	4264	61.5	1277	4	US-09-645-791-12
23	4249.5	61.3	1298	4	US-09-645-456A-14
24	4249.5	61.3	1298	4	US-09-425-324A-14
25	4249.5	61.3	1298	4	US-09-645-791-14
26	4239.5	61.2	1306	4	US-09-645-456A-10
27	4239.5	61.2	1306	4	US-09-425-324A-10

28	4239.5	61.2	1306	4	US-09-645-791-10	Sequence 10, Appl
29	4195.5	60.5	1297	4	US-09-688-188B-14	Sequence 14, Appl
30	4195.5	60.5	1297	4	US-09-291-417D-14	Sequence 14, Appl
31	4021.5	58.0	1239	4	US-09-688-188B-13	Sequence 13, Appl
32	4021.5	58.0	1239	4	US-09-291-417D-13	Sequence 13, Appl
33	3951.5	57.0	1233	4	US-09-688-188B-89	Sequence 89, Appl
34	3951.5	57.0	1233	4	US-09-291-417D-89	Sequence 89, Appl
35	3886.5	56.1	1233	4	US-09-645-456A-35	Sequence 35, Appl
36	3886.5	56.1	1233	4	US-09-425-324A-35	Sequence 35, Appl
37	3886.5	56.1	1233	4	US-09-645-791-35	Sequence 35, Appl
38	2708.5	39.1	1109	4	US-09-688-188B-88	Sequence 88, Appl
39	2708.5	39.1	1109	4	US-09-291-417D-88	Sequence 88, Appl
40	1862.5	26.9	1227	4	US-09-688-188B-105	Sequence 105, App
41	1862.5	26.9	1227	4	US-09-291-417D-105	Sequence 105, App
42	1457	21.0	275	4	US-09-579-664B-8	Sequence 8, Appl
43	1309	18.9	444	3	US-09-221-235-14	Sequence 14, Appl
44	1309	18.9	444	3	US-09-221-928-14	Sequence 14, Appl
45	1309	18.9	444	3	US-09-221-527-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-09-688-188B-15
; Sequence 15, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 039602/0328
; CURRENT APPLICATION NUMBER: US/09/688.188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-15

Query Match	96.1%	Score 6659;	DB 4;	Length 1326;
Best Local Similarity	95.8%	Pred. No. 0;		
Matches 1270;	Conservative	3;	Mismatches	3;
Indels	50;	Gaps	3;	
QY	36	YGVYKGRHVKTGQLAAIKVMDVTEDEEEIKQINMLKKYSHRNIAITYYGAFIKKSP 95		
DB	2	FGEVYEGRHVKTGQLAAIKVMDVTEDEEEIKQINMLKKYSHRNIAITYYGAFIKKSP 61		
QY	96	GNDQLMLVMEFCGAGSVTDLVKNTKGNALKEDECIAYICREILRGLAHLHAHKVHRDIK 155		
DB	62	GNDQLMLVMEFCGAGSVTDLVKNTKGNALKEDECIAYICREILRGLAHLHAHKVHRDIK 121		
QY	156	QNVLLTENAEVKLVDFGVSAQLDRTVGRNRNFTGTPTVWMAPEVTACDENPDATYDVRSD 215		
DB	122	QNVLLTENAEVKLVDFGVSAQLDRTVGRNRNFTGTPTVWMAPEVTACDENPDATYDVRSD 181		
QY	216	IWSLGITAIEMAEAPPLCDMHPRALFLIPRNPPLPKSKWKKKFDIDTCLIKTYL 275		
DB	192	IWSLGITAIEMAEAPPLCDMHPRALFLIPRNPPLPKSKWKKKFDIDTCLIKTYL 241		
QY	276	SRPTEQLLKPPFIRDOPTERQVRIQLKDHIDRSKKRKEKEETEYVSGSEEDDSHGE 335		
DB	242	SRPTEQLLKPPFIRDOPTERQVRIQLKDHIDRSKKRKEKEETEYVSGSEEDDSHGE 301		
QY	336	EGEPSSINNVGSESTLRREFLRLOENKNSSEALKQQOOLQQOORDPEAHKHLHORQ 395		


```
Db 722 GPNASSNPDLRRSPGERSDVLPPASHGHLPPQAGSLRRNRVGVSSKPDSSPVLSPGNK 781
Qy 767 AKPDHRRSPGADPVLKERTLEAPRPPKAMDYSSSEVESSDDDEEGGGPAE 826
Db 782 AKPDHRRSPGADPVLKERTLEAPRPPKAMDYSSSEVESSDDDEEGGGPAE 841
Qy 827 GSRDTPGSGDGTDSVSTVMVHDVVEITGTQPPYGGGTVMVQRTPEERNLLHADSNY 886
Db 842 GSRDTPGGR-DGDTDSVSTVMVHDVVEITGTQPPYGGGTVMVQRTPEERNLLHADSNY 900
Qy 887 TNLDPVOPSHSPSNSKQSPSSKDGSGDVQSRGLVAPKSSFTMFVDLGIYOPGSG 946
Db 901 TNLDPVOPSHSPSNSKQSPSSKDGSGDVQSRGLVAPKSSFTMFVDLGIYOPGSG 960
Qy 947 DSIPITLVGGEGTRLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKGFNSILCAA 1006
Db 961 DSIPITLVGGEGTRLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKGFNSILCAA 1020
Qy 1007 LGVNVLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKLRY 1066
Db 1021 LGVNVLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKLRY 1080
Qy 1067 LSWLRNKILHNDPEVEKKGWTTVGDMEGCGHYRVVYERIKFLVIALKSSVEVYAWAPK 1126
Db 1081 LSWLRNKILHNDPEVEKKGWTTVGDMEGCGHYRVVYERIKFLVIALKSSVEVYAWAPK 1140
Qy 1127 PYHKFMAFKSPADLPHRPLLDVLTVEEGORLKVYIGSSAGFHAVDVSNGSYDIYPVHI 1186
Db 1141 PYHKFMAFKSPADLPHRPLLDVLTVEEGORLKVYIGSSAGFHAVDVSNGSYDIYPVHI 1200
Qy 1187 QSQITPHAIIFLNTDGMELLCYDEGVVNTYGRIIKDVVLQWGEPMPTSVAYICSNOI 1246
Db 1201 QSQITPHAIIFLNTDGMELLCYDEGVVNTYGRIIKDVVLQWGEPMPTSVAYICSNOI 1260
Qy 1247 MGWGEKAIEIRSVETGHLDDGVFMHKAORLKLFCERNDKVPFASVRSGSSQVYFMTLNR 1306
Db 1261 MGWGEKAIEIRSVETGHLDDGVFMHKAORLKLFCERNDKVPFASVRSGSSQVYFMTLNR 1320
Qy 1307 NCIMNW 1312
Db 1321 NRIMNW 1326

RESULT 3
US-09-645-456A-13
; Sequence 13, Application US/09645456A
; Patent No. 6562580
; GENERAL INFORMATION:
; APPLICANT: FU, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND METHODS OF USE
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,456A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US/09/425,324
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-456A-13

Query Match 63.1%; Score 4370.5; DB 4; Length 1324;
Best Local Similarity 64.7%; Pred. No. 7.8e-258;
Matches 899; Conservative 137; Mismatches 210; Indels 143; Gaps 31;

Qy 1 MGDPAARSLLDDIDLSALRDPAGIFELVELVGNCTYGVYKGRHVKTGQLAAIKVMDVTG 60
Db 61 DEEBEIKOENMLKKYSHRNIIATYYGAFIKKSPGNDQLWLWNEFCGAGSVDTLVKQT 120
Db 61 DEEBEIKOENMLKKYSHRNIIATYYGAFIKKSPGNDQLWLWNEFCGAGSVDTLVKQT 120
Qy 121 KGNALKEDCAYICREITLGLAHLHAKVTHRDIKGONVLLTENAELVKLVDFGVSAQLDR 180
Db 121 KGNLTKEEWAIYICREITLGLSHLHQHKVTHRDIKGONVLLTENAELVKLVDFGVSAQLDR 180
Qy 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDISLGLITAJEMAEGAPPLCDMHPMR 240
Db 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDISLGLITAJEMAEGAPPLCDMHPMR 240
Qy 241 ALFLIPNPRLPKSKWKKKFDIDTCLIKTVLSRPPTQLKFPFIRDOPTERQVRI 300
Db 241 ALFLIPNPAPLKKSKKKKQSFIESCLVKNHSQPATEQLMKHPFIRDOPTERQVRI 300
Qy 301 QLKDHIDRSRKKEETEVEYSGSEEDDSHGEPESSIMNVPGESTLURRFLRLQQ 360
Db 301 QLKDHIDRTKKRGEKDETEYSGSEEBEEN-DSGEPSILNPGESTLURRDFLRQL 359
Qy 361 ENKNSBALKQOQOQOQOQDPEAHIKHLHQRRIIEQKEERRRVEEQRREREQRK 420
Db 360 ANKERSEALRRQ--LEQOQRENEEHKQLLAERQKIEQKEQRRLEEQRREREQRK 417
Qy 421 LQEKQORRLLEDQALREERERQAEQBYKRLQLEEQSERLQRLQQLQOEHLAYLSLQ 480
Db 418 QOERQRRHYBEQ--MRREERRRAEHEQBYKRLQLEEQSERLQRLQQLQOERDYLVSQ 475
Qy 481 QOQOQOQLQOQOQOQLPGDRKPLHYGRGNVPADKAWAREVEERTMKNQKNSPLAKS 540
Db 476 HQREQR--PVEKKPLYHYKBSGSPSEKPAWAKEVEERSRLNR-QSSPAMPH 524
Qy 541 KPGSTGPEPTPOQASPG-----PPGPLSOTPPMQRPVEPO-----EGP-----HKS 581
Db 525 KVANRISDPNLPPESEFSISGVQP---ARTPPMLRPVDPQIPHVLVAKSGPALTASQS 581
Qy 582 LQDQPTNLAAFPASHDPDPAIPAPTAPTSARGAVIRONSDDPTSEGPGPSP-----NPP 635
Db 582 VHEQPTKGLSGFQEALN-----VTSHRVEMPRQNSDPTSENP-PLPTRIEKFDRS 630
Qy 636 AWVRPDNEAPKVPQRTSSITATALTSGAGSRPAQAVRAPRPSNSAWQIYLQRAERGT 695
Db 631 SWLRQEEIDIPKVPQRTTSSIPALARKNSPGNSALG-----PRIGS----- 672
Qy 696 PKPPGPPAQPAPPNASSNPDLRRSDPGWE-----RSDSVLPASH-----GHLP- 739
Db 673 -----QPIRASNPDLRTEPILSPLORTSSGSSSSSTPSSQPSOGGSPG 720
Qy 740 -QAGSLERNRVGASSKLDSPVL--SPGNKAKPDHR--SRGPRPADFVL-----LKERTL 790
Db 721 SQAGSSERTVRANSKSESGPVLPHPEA-KVKPEESRDITRPSRPADILTALAKELRELRI 779
Qy 791 DEAPRPPKAMDYSSSEVESSDDDEEGGGPAES---RDTP-----GGRSDGDTDS 842
Db 780 EETNRPMKKVTDYSSSESESESESESESETHDGTAVSDIPRIIPTGAPSGNEOYN 839
Qy 843 VSTMVVDHVEITGTQPPYGG-----GTMVQVQRTPEERNLLHADSNYGT---NLPDVVQ 894
Db 840 VGMVGTGLE--TSHADSPSGSISREGTLMIRETSGEKKRSGHSDNSGFAGHINLPDLVQ 897
Qy 895 PSHS-----PTNSKGGQPPSKO-GSGDYQSRGLVAFKGSFTWFDLGIYQPGS----- 945
Db 898 QSHSPAGTPTTEGLGRVSTHSDMDSGTEYGMG---SSTKASFTPFVDPRVYQTSPTDEDE 954
Qy 946 -GDSIPITLVGGEGTRLDQLQY-DVRKGSVVNVNPTNTRAHSETPEIRKYKGFNSIL 1003
Db 955 EDESSAALFTSELLRQEQAKLNEARKISVVNVNPTNIRPHSDTPEIRKYKGFNSIL 1014
Qy 1004 CAALGWNVLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKL 1063
Db 1015 CAALGWNVLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKL 1074
```



```
Db      840  SETHDGTAVSDIPRLIPTGAGSNEQYVGMVGTGLE--TSHADSFSGSISREGTLM 897
Qy      869  QRTPEERNLHADSNYGT---NLPDVQPSHS-----PTENSKGQSPPSKD--GSGDYQSR 920
Db      898  RETSGEKRSCHSDSNGFAGHINLPDLVQOQSHSPAGTPTGELGRVSTHSEMDSGTEYGM 957
Qy      921  GLVKAPGKSSFTMFVDLGIYQPGGS-----GDSIPITALVGGEGTDLQIY--DVRKGSV 974
Db      958  G---SSTKASFTFPVDPVRYQTSPTDEDEBESSAAALFTSELLRQEQAKLNEARKISV 1014
Qy      975  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGKVVYL 1034
Db      1015  VNVNPTNIRPHSDTPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGKVVYL 1074
Qy      1035  IGRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDM 1094
Db      1075  INRRFQOMDVLGLNLLVGTISGKRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDM 1134
Qy      1095  GCGHYRVVKYERIKFVLVIALKSVSEVYAWAPKPYHKFMFKFADLPKRPILLVLTVE 1154
Db      1135  GCIHVKKVYERIKFVLVIALKNAVEIYAWAPKPYHKFMFKFADLPKRPILLVLTVE 1194
Qy      1155  QRLKVIYSSAGFHAVDVDSGNSDIYIPVHIQSOITTHAIIFLNTDGMELLCYDEBG 1214
Db      1195  QRLKVIYSSAGFHAVDVDSGNSDIYIPVHIQSOITTHAIIFLNTDGMELLCYDEBG 1254
Qy      1215  VYVNTYGRITKDVVLQWGMPTSVAYICNSQINQIMGGEKAIEIRSVETGHLGDFV 1274
Db      1255  VYVNTYGRITKDVVLQWGMPTSVAYIHSNQINQIMGGEKAIEIRSVETGHLGDFV 1314
Qy      1275  RLKFLCERNDKVFFASVRSGSSQVYFMTLNRNCIMNW 1312
Db      1315  RLKFLCERNDKVFFASVRSGSSQVYFMTLNRNSMWN 1352

RESULT 10
US-09-425-324A-11
; Sequence 11, Application US/09425324A
; Patent No. 6562591
; GENERAL INFORMATION:
; APPLICANT: FU, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AN
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DRR
; CURRENT APPLICATION NUMBER: US/09/425,324A
; CURRENT FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-425-324A-11

Query Match 62.7%; Score 4346; DB 4; Length 1353;
Best Local Similarity 63.4%; Pred. No. 2.5e-256;
Matches 899; Conservative 137; Mismatches 210; Indels 172; Gaps 32;

Qy      1  MGDPAARSLDDIDLSALRDPAGIFELVEVWGNGTYGYQYGRHVKTGQLAAIKVMDVTE 60
Db      1  MASDSPARSLDEIDLALRDPAGIFELVELVNGTYGYQYGRHVKTGQLAAIKVMDVTE 60
Qy      61  DEEEETKQINMLKKYSHRNATYVYGAFTKSPQNDOLWLVMEFCGAGSVTLVKN 120
Db      61  DEEEETKQINMLKKYSHRNATYVYGAFTKKNPPGMDQLWLVMEFCGAGSVTLIKNT 120
Qy      121  KGNALKEDCIATYICREILAGLAHLAHKVIHRDIKQGNVLLTENAELVDFGVSAQLDR 180
Db      121  KGNLKEEWIATYICREILAGLSHLHQHKVIHRDIKQGNVLLTENAELVDFGVSAQLDR 180
```

```
Qy      181  TVGRENFTIGPYMMAPEVIACDENPDATYDYSIDWSLIGITATEMAAGAPPLCDHMPMR 240
Db      181  TVGRENFTIGPYMMAPEVIACDENPDATYDFKSDLSWLSLIGITATEMAAGAPPLCDHMPMR 240
Qy      241  ALFLIPRNPRLKSKWKKFIDFIDTCLIKTYLSREPTBQLKFPFIRDOPTERQVRI 300
Db      241  ALFLIPRNPRLKSKWKKFQSFIESCLVKNHSORPATBQLMKHPFIRDOPTERQVRI 300
Qy      301  QLKDHIDRSKRKEKETETEYSGSEBDDSHGEEGEPSSIMVPGESTLRREFLRLOO 360
Db      301  QLKDHIDRTKKRGEKDETEYSGSEBEEEN--DSGEPSSILNPGESTLRREFLRLOO 359
Qy      361  ENKNSSEALQKQOQLQQOQOORDPEAHIKHLAHRORRLEEOKEERRRVEEQORRREROK 420
Db      360  ANKERSEALRQO--LEOQORENEEHKQQLAERQKRIEEOKEORRUEEQORRREROK 417
Qy      421  LQKEEQOQRLDMQALRREERQAREQ-----QREERERRAEHEQYIRQLBEEQRLQQLLHEQALL 451
Db      418  QQEREQRHVEEQ--MRREERERRAEHEQYIRQLBEEQRLQQLLHEQALL 475
Qy      452  KRQLEEQORQERLORLOQOEHAYLKSLOOQOQOQOQOQOQOQOQOQOQOQOQOQO 511
Db      476  KRQLEEQORQERLORLOQOERDYLVSQHQOQOQOQOQOQOQOQOQOQOQOQOQO 525
Qy      512  NPADKPAWAEVREERTMKNQONSPLAKSPGSGTGPPEPIPAQSPG-----PGPLSOT 565
Db      526  SPSEKPAWAEVREERSRLNR--QSSPAMPKHVANISDNLNPPRSSEFSISGVQP---ART 581
Qy      566  PPMORPVEPQ-----EGP-----HKSLOQOQPTNLAAFPASHPDPAIPAPTATPSA 612
Db      582  PPMLRPVDVQPLPHLVAVKSGQPALTASQSVHEQPTKGLSGFQALN-----VTSH 631
Qy      613  RGAVIRQNSDPTSGPGSP-----NPPAVWRPDNEAPPKVPQRTSIAIATNTSGAG 666
Db      632  RVEMPRQNSDPTSENP--PLPTRIEKFDRSSMLRQEBDIPPKVPQRTTTSIPALARKSPG 690
Qy      667  SRPAQAVPARPRSNSAQIYLQRAERCTPKPPGPPAPPPNASSNPDLRRSDPGWE- 725
Db      691  NGSALG-----PRLGS-----QPTRASHNPDLRRTEPILES 720
Qy      726  -----RSDSVLPASH-----GHLP--QAGSLERNRVGASSKLDSSPVL--SPGNKA 767
Db      721  PLQRTSSGSSSSSTPSSQPSQSGSQSGQAGSSERTVRANSKSESPVL.PHEPA-KV 779
Qy      768  KPDDHR--SRPGRPADFVL-----LKERTLDRAPPPKAMDYSSSESESESESEDEE 821
Db      780  KPESRDITRPSRPADLTALAKELRELRIEETNRPMKKVTDYSSSESESESESEDEE 839
Qy      822  GGAEGS---RDTG-----GGRSDGDTDSVSTMVVDVEITGTQPPYGG-----GTWV 868
Db      840  SETHDGTAVASDIPRLIPTGAGSNEQYVGMVGTGLE--TSHADSFSGSISREGTLM 897
Qy      869  QRTPEERNLHADSNYGT---NLPDVQPSHS-----PTENSKGQSPPSKD--GSGDYQSR 920
Db      898  RETSGEKRSCHSDSNGFAGHINLPDLVQOQSHSPAGTPTGELGRVSTHSEMDSGTEYGM 957
Qy      921  GLVKAPGKSSFTMFVDLGIYQPGGS-----GDSIPITALVGGEGTDLQIY--DVRKGSV 974
Db      958  G---SSTKASFTFPVDPVRYQTSPTDEDEBESSAAALFTSELLRQEQAKLNEARKISV 1014
Qy      975  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGKVVYL 1034
Db      1015  VNVNPTNIRPHSDTPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGKVVYL 1074
Qy      1035  IGRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDM 1094
Db      1075  INRRFQOMDVLGLNLLVGTISGKRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDM 1134
Qy      1095  GCGHYRVVKYERIKFVLVIALKSVSEVYAWAPKPYHKFMFKFADLPKRPILLVLTVE 1154
Db      1135  GCIHVKKVYERIKFVLVIALKNAVEIYAWAPKPYHKFMFKFADLPKRPILLVLTVE 1194
Qy      1155  QRLKVIYSSAGFHAVDVDSGNSDIYIPVHIQSOITTHAIIFLNTDGMELLCYDEBG 1214
```

Db 1195 QRLKVFSGHTGPHVTDVDSGNSYDIYIPSHIQGNITPHAIIVLPKTDGMEMLVCYDESG 1254
Qy 1215 VYVNTYGRILIKDVLQWGMPTSVAYICSNQIMGGEKAIEIRSVETGHLGDFVFMHRAQ 1274
Db 1255 VYVNTYGRITKDVVLQWGMPTSVAYIHSNQIMGGEKAIEIRSVETGHLGDFVFMHRAQ 1314
Qy 1275 RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
Db 1315 RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1352

RESULT 11
US-09-645-791-11
; Sequence 11, Application US/09645791
; Patent No. 6569658
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Fu, Alan C
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS; COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: A-68344-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,791
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 09/425,324
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-791-11

Query Match 62.7%; Score 4346; DB 4; Length 1353;
Best Local Similarity 63.4%; Pred. No. 2.5e-256;
Matches 899; Conserved 137; Mismatches 210; Indels 172; Gaps 32;

Qy 1 MGDPAARSLLDIDLSALRDPAGIFELVELVWNGTYGYQVYKGRHVKTGQLAAIKVMDVTE 60
Db 1 MASDSPARSLLDIDLSALRDPAGIFELVELVWNGTYGYQVYKGRHVKTGQLAAIKVMDVTG 60
Qy 61 DEEEELKQBINLKKYSHRNATYGYAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120
Db 61 DEEEELKQBINLKKYSHRNATYGYAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120
Qy 121 KGNALKEDCIAYICREILRGLAHLHAHKVHHRDIKQNVLLTENAEEVKLVDFGSAQLDR 180
Db 121 KGNLKEEWIAYICREILRGLSHLQHKVHHRDIKQNVLLTENAEEVKLVDFGSAQLDR 180
Qy 181 TVGRNTFTGTYWPAVEVIACENPDATYDYSRDIWSLIGITAIEMAEAGAPCLDMHMR 240
Db 181 TVGRNTFTGTYWPAVEVIACENPDATYDYSRDIWSLIGITAIEMAEAGAPCLDMHMR 240
Qy 241 ALFLIPRNPPRLKSKKSKKFIIDTCLIKTILSRPTEOLLKFPTRDOPTEROVRI 300
Db 241 ALFLIPRNPPRLKSKKSKKFIIDTCLIKTILSRPTEOLLKFPTRDOPTEROVRI 300
Qy 301 QLKDHIDRSKRKRGEKEETEYYSGESEEDDSHGEGEPSSIMNVPGESTLRRREFLRQQ 360
Db 301 QLKDHIDRTKKRGEKDETEYYSGESEEBEEN-DSGEPSSILNLPGESTLRRDFLRQL 359
Qy 361 ENKNSSEALKQOOQLOOQOORPEAHIKHLLHQRRRIEEOKEERRRVEEQORRREORK 420
Db 360 ANKERSEALRQQ--LEOQORENEEHKQQLLAEROKRIEEOKEORRRIEEOQORREKLRK 417
Qy 421 LOEKEORLEDMQALRRREERROAREQO-----EY 451
Db 418 QOERQORRYEQ--WREERERRABHEQYIRROLEEEQROLETLQOQLLHEQALLLEY 475

Qy 452 KRKLEEQORSERLQRLQOQEHAYLKSLOQOQOQOQLOKQOQOQLLPGDKRPLVHYGRGM 511
Db 476 KRKLEEQORAEQLQRLQKQERDYLVLQHQEQOR-----PVEKKPLVHYKEGM 525
Qy 512 NPADKPAAWREVERTRANKOONSPLAKSKGSGTGPPEPIPOASPG-----PGPLSQT 565
Db 526 SPSEKPAWAKEVERSRNLNR-QSSPAMPKHVANIISDPNLPRSESFISGVQP---ART 581
Qy 566 PMORPVEPQ-----EGP-----HKSLODQPTENLAAPPASHDDPPAIPAPTATPSA 612
Db 582 PMLRPVDPQIPLHAVKSGQPALITASQVHEQTKGLSGFQEALN-----VTSH 631
Qy 613 RGAIVRQNSDPTSEGPSP-----NPPAVWRPDNEAPPKVPQRTSSIALNTSGAGG 666
Db 632 RVEMPRQNSDPTSENP-PLPTRIBKFDORSMLRQEDIPPKVPQRTTSSIPALARKNSPG 690
Qy 667 SRPAQAVRPRNSAWIYLORRAERCTPKPPPPAQPQPPPPNASSNPDLRRSDPGWE- 725
Db 691 NGSALG-----PRLGS-----QPIRASNPDLRTEPILES 720
Qy 726 -----RDSVLPASH-----GHLP--QAGSLERNRVGASSKLDSSPVL--SPGNKA 767
Db 721 PLQRTSSGSSSSSTPSSQSGSQSGQSGSERTRVRANSKSESPVLPEPA-KV 779
Qy 768 KPDDHR--SRFGRPADFVL-----LKERTLDRAPRPKKAMDYSSSESSEESDEDEGE 821
Db 780 KPESRDITRSPADLTALAKELRELRIETNRPMKKVTDYSSSESSESESESEDEGE 839
Qy 822 GGPAGS---RDTP-----GGRSDGDTDSVSTVWVHVEEITGTQPPYGG-----GTWV 868
Db 840 SETHDGTAVASDIPRLIPTGAPSGNEQYVNMVGVTGLE--TSHADSPSGSISREGTLM 897
Qy 869 QRTPEERNLHADSGNYT---NLPDVVQPSHS---PTENSKGQSPSPK-D-GSDYQSR 920
Db 898 RETSGEKRSCHSDSNGFAGHINLPDLVQOQSHSPAGTPTTEGLGRVSTHSQEMDSGTEYGM 957
Qy 921 GLVKAPGKSFMTFVDLGIYQPGS-----GDSIPITALVGGEGTRLDQLQY-DVRKGSV 974
Db 958 G---SSTKASFTFPVDPVRYQTSPTDEDEDEESSAAALFTSELLRQEQAKLNEARKISV 1014
Qy 975 VNVPTNTRAHSETPETIRKYKRNSEILCAALGVNLLVGTENGLMLLDRSGGQKVYGL 1034
Db 1015 VNVPTNIRPHSDTPETIRKYKRNSEILCAALGVNLLVGTENGLMLLDRSGGQKVYNL 1074
Qy 1035 IGRRFQOMDVLGILLITISGKRNKLVYVYLSWLNKILHNDPVEVEKKGQWTVGDME 1094
Db 1075 INRRRFQOMDVLGSLNVLVTISGKKNKLVYVYLSWLNRLHNDPVEVEKKGQWTVGDLE 1134
Qy 1095 GCGHYRVVKYERIKFLVIALKSVVEVYAWAPKPYHKFMAFKSPADLPHRPLLVLTVEEG 1154
Db 1135 GCIHYKVKYERIKFLVIALKNAVEIYAWAPKPYHKFMAFKSPADLQHKPLLVLTVEEG 1194
Qy 1155 QRLKVIYSSAGFHAVDVDSGNSYDIYIPVHIQSOITPHAIIFLPTDGMELICYDEEG 1214
Db 1195 QRLKVIYSGHTGFHVIDVDSGNSYDIYIPSHIQGNITPHAIIVLPKTDGMEMLVCYDEEG 1254
Qy 1215 VYVNTYGRILIKDVLQWGMPTSVAYICSNQIMGGEKAIEIRSVETGHLGDFVFMHRAQ 1274
Db 1255 VYVNTYGRITKDVVLQWGMPTSVAYIHSNQIMGGEKAIEIRSVETGHLGDFVFMHRAQ 1314
Qy 1275 RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
Db 1315 RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1352

RESULT 12
US-09-393-569-2
; Sequence 2, Application US/09393569
; Patent No. 6277979
; GENERAL INFORMATION:
; APPLICANT: BINGHAM, SHARON
; APPLICANT: CASE, PATRICK
; APPLICANT: LAWSON, SALLY NEALE

Best Local Similarity 63.0%; Pred. No. 1e-255;	
Matches 898; Conservative 138; Mismatches 210; Indels 180; Gaps 32;	
QY	1 MGDPAARSDDIDLSALRDPAGIFELVEVNGNTGYQVYKGRHVKTQQLAAIKMVDYTE 60
DB	1 MASDSPARSLDEIDLSALRDPAGIFELVELVNGNTGYQVYKGRHVKTQQLAAIKMVDVG 60
QY	61 DEEEI1KQEI1NMLKKYSHHRI1ATYYGAFIKKSPGNDQDLVLMVEFCAGSVTDLVKNT 120
DB	61 DEEEI1KQEI1NMLKKYSHHRI1ATYYGAFIKKPPGMDQDLVLMVEFCAGSVTDLKNT 120
QY	121 KGNALKEDCTAYICRETLRGLAHLHAHKVTHRDIKQNVLLTENAENVLDPFGVSAQLDR 180
DB	121 KGNLTKEWYAYICRETLRGLSHLQHVTHRDIKQNVLLTENAENVLDPFGVSAQLDR 180
QY	181 TVGRNRTFIGTPYMAPEVIACDENPDATVYRSDIWSLGI1TAIEMASGAPLDCMHPMR 240
DB	181 TVGRNRTFIGTPYMAPEVIACDENPDATVDFKSDLWSLGI1TAIEMASGAPLDCMHPMR 240
QY	241 ALFLIPRNPPLKSKWKKFIDFIDTCLIKTYLSRPTEQLLKFPPIRQDPTERQVRI 300
DB	241 ALFLIPRNPRLKSKWKKFQSFIESCLVKNHSQRPAEQLMKHPPIRQDPNERQVRI 300
QY	301 QLKDHIDRSKRGEKETEYIYSGSEEDDSDHGEGBESSIMVNVPGESTLRRLFLKQQ 360
DB	301 QLKDHIDRTKKGGEKDETEYIYSGSEEBEEN--DSGEPSSILNIPGSESTLRRLFLQL 359
QY	361 ENKNSBALQKQOQLOQOQOORDEPAHMKLLHQRORRIEEOKEERRRVEEQRREBORK 420
DB	360 ANKERSALRQO--LQQQRENEEHKQLLAERKRIEEOKEQRRLVEEQORKELEURK 417
QY	421 LQXEQOORLMDQALREERERQAREQ-----EY 451
DB	418 QQEREQRRHYEQ--MEREERERRAHEQYIRRLQLEEEQRQLEILQOQLLHEQALLLEY 475
QY	452 KRKOLEHQRQSERLQLOQOEHA1YKLSLQOQOQOQLOKQOQOQLLPGDRPLHYHGRGM 511
DB	476 KRKOLEQROAERLQRLKOERDYLVSLOHQEQR-----PVEKKPLHYHEGM 525
QY	512 NPADKPAWAVEERTMNMKQONSPLAKSKPGSTGPEPPIQOASPG-----PGPLSOT 565
DB	526 SPSEKPAWAVEERSKLNR--QSSPAMPKHVAVNRISDNLPRSESPISGVQP---ART 581
QY	566 PPMQRPVEPQ-----EGP-----HKSLQDQPTRNAAPFASHDDPDPAIPAPTATPSA 612
DB	582 PPMRLRPVDQI1PHLVAKVSGPAL1TASQSVHEQPTKGLSGFQALN-----VTSH 631
QY	613 RGA1VRONSPTSEGPGPSP-----NPAAWVRPDNEAPPKVQPTQSS1ATALNTSGAGG 666
DB	632 RVENPRQNSDPTSENP--PLPTRIEKDFDRSWSLRBEDIPPKVPQRTTIS1S1PALARKNSPG 690
QY	667 SRPAQAVRAPRNSAWQIYLORRAERGTPKPPGPPAPGPPNASSNPDLRRSDPGWE--725
DB	691 NGSALG-----PRLGS-----QPTRASSNPDLRRPTILES 720
QY	726 -----RSDSVLPASH-----GHLF--QAGSLERNRVGASSKLDSDPVL---SPGNKA 767
DB	721 PLQRTSSGSSSSSTSPSSQSGSQPSQAGSSERTFVANSKESGSPVLPEHPA-KV 779
QY	768 KPDDHR--SRGPRADPVL-----LKERTLEAPRPKKAMDYSSSSSEEVSS 813
DB	780 KPESRODITRPSRPASYKKAIDED1TALAKELRELRIETNRPMKKVTDYSSSSSEESS 839
QY	814 EDDEEEGEGGPAEGS---RDTP-----GGRSDGDDTSVSTMVHVDVEITGTQPYCG--863
DB	840 EEEEDGESETHDGTVAVSDIPLRI1PTGAPGSENGQNVGMVGTGHLG--TSHADSFSGSI 897
QY	864 ---GTWVVQRTPEERNLHADNSGYT---NLPDVVVQPSHS-----PTENSKGQSPSKD- 912
DB	898 SRECTLMIRETSGEKKSGHSDNSGFGACHINLPDLVQOSHSPAGTPTGELGFEVTHSQEM 957
QY	913 GSGDYQSRGLVLVAPGKSSFTFVLDLGIYQPGGS-----GDSIPITALVGGSEGLDQLQY 967

Db	958	DSGTEYGMG---SSTKASFTFPFVDPVRYQTSPTDDEDEDESSAAALFTSELLRQEQAKL	1014
Qy	968	-DVRKGSVVNVNPTNTRAHSETPEIRKYKRFNSEILCAALMGVNLIVGTENGLMLLDRS	1026
Db	1015	NEARKLSVVNVNPTNIRPHSDTPEIRKYKRFNSEILCAALMGVNLIVGTENGLMLLDRS	1074
Qy	1027	GQSKVYGLIGRRPFQMDVLVEGLNLIITISGKKNLURVYVLSWLRNKILHNDPEVEKKQG	1086
Db	1075	GQSKVYNLIIINRRRFQMDVLEGLNVLVTISGKKNLURVYVLSWLRNRIHNDPEVEKKQG	1134
Qy	1087	WTTVTGDMGCGHYRVVVKYERIKFLVIALKSSVEVYAWAKPKYHKFWAFKSFADLPHRPLL	1146
Db	1135	WTTVTGDELEGCIHYKVKYERIKFLVIALKNAVEIYAWAKPKYHKFWAFKSFADLQHKPLL	1194
Qy	1147	VDLTVBEGQRLKVIYSSAGFAVDVDSGNSYDIYIPVHIQSOITPHAIIFLPNTDGMEM	1206
Db	1195	VDLTVBEGQRLKVIYSGHTGFHVIDVDSGNSYDIYIPSHIQGNITPHAVILPKTDGMEM	1254
Qy	1207	LLCYEDGEGVYVNTYGRIIKDWVLQWGMPTSVAYICSNQIMGWGEKAIIRSVETGHLDG	1266
Db	1255	LVCYEDGEGVYVNTYGRITKDWVLQWGMPTSVAYIHSNQIMGWGEKAIIRSVETGHLDG	1314
Qy	1267	VFMHKAQRLKFLCERNDKVPFASVRSGGSSQVVFMTLNRNCIMNW	1312
Db	1315	VFMHKAQRLKFLCERNDKVPFASVRSGGSSQVVFMTLNRNSMNV	1360
RESULT 14			
US-09-645-456A-34			
; Sequence 34, Application US/09645456A			
; Patent No. 6562580			
; GENERAL INFORMATION:			
; APPLICANT: Fu, C. Alan			
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOS			
; FILE REFERENCE: METHODS OF USE			
; CURRENT APPLICATION NUMBER: US/09/645,456A			
; CURRENT FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: US/09/425,324			
; PRIOR FILING DATE: 1999-10-21			
; NUMBER OF SEQ ID NOS: 36			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 34			
; LENGTH: 1360			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: synthetic			
US-09-645-456A-34			

Query Match	62.6%;	Score 4336;	DB 4;	Length 1360;			
Best Local Similarity	63.0%;	Pred. No. 1e-255;					
Matches	899;	Conservative 138;	Mismatches 210;	Indels 180;	Gaps 32;		
Qy	1	MGDPAPARSLDDIDL	SALRDPAGIFELV	EVWNGTYGYQVYKGRHVKTGOLAAIKWMDVTE	60		
		:	:	:			
Db	1	MASDSPARSIDEIDL	SALRDPAGIFELV	ELVNGTYGYQVYKGRHVKTGOLAAIKWMDVTG	60		
Qy	61	DEEEETKQETNMLKKY	SHRNTATYYGATIK	SPGNDQDLWLWBEFCGAGSVTDLAKVT	120		
		:	:	:			
Db	61	DEEEETKQETNMLKKY	SHRNTATYYGATIK	PGMDQDLWLWBEFCGAGSVTDLAKVT	120		
Qy	121	KGNALKEDCIATY	CREILRGLAHLAHK	VIHRDIKQNVLLTENAEBVKLVDFGVSQALOR	180		
		:	:	:			
Db	121	KGNTLKEWIAVY	CREILRGLSHLHQ	KVIHRDIKQNVLLTENAEBVKLVDFGVSQALOR	180		
Qy	181	TVGRNRTFTIGTPY	WMAPEVIACDENP	DATYDRSDIWSLGI	TAIEMAEAGAPPLCDMHPMR	240	
		:	:	:			
Db	181	TVGRNRTFTIGTPY	WMAPEVIACDENP	DATYDFKSDLWSLGI	TAIEMAEAGAPPLCDMHPMR	240	
Qy	241	ALFLIPRNPPRLK	SKWSKKFIDT	CTLIKTYYSRPPTE	OLLKPFPIRDPQTRQVRI	300	
		:	:	:			
Db	241	ALFLIPRNAPRLK	SKWSKKFOSFT	IESCLVNKHSORP	ATEOLMKHPFI	TRDPNBRQVRI	300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:43:18 ; Search time 154 Seconds
(without alignments)
2407.158 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAAPSLDIDLSALRD.....SGSSQVYFMTLNRNCINW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6929	100.0	1312	7 ADE34150	Human mis
2	6660.5	96.1	1295	5 AAO18508	Human ins
3	6659	96.1	1326	2 AAY55933	Human ZC3
4	6646.5	95.9	1303	4 AAM79153	Human pro
5	6646.5	95.9	1303	5 AAO18507	Human ins
6	6628.5	95.7	1303	4 AAO18507	Human nov
7	6622	95.6	1332	4 AAE10612	Human nov
8	6603.5	95.3	1276	7 ADE34154	Human mis
9	6311	91.1	1244	4 AAE10614	Human nov
10	6304.5	91.0	1273	4 AAE10613	Human nov
11	5651	81.6	1701	4 ABG24020	Novel hum
12	4370.5	63.1	1324	4 AAB68221	Amino aci
13	4360.5	62.9	1332	4 AAB68217	Amino aci
14	4346	62.7	1353	4 AAB68219	Amino aci
15	4336	62.6	1360	3 AAB85263	Human pro
16	4336	62.6	1360	4 AAB50059	Large NIK
17	4336	62.6	1360	5 AAO18506	Human ins
18	4336	62.6	1385	7 ADP14162	Human sfc
19	4325	62.4	1385	4 AAM79405	Human pro
20	4274	61.7	1268	7 ADE31765	Human pro
21	4274	61.7	1269	4 AAB68223	Amino aci
22	4264	61.5	1277	4 AAB68220	Amino aci
23	4249.5	61.3	1298	4 AAB68222	Amino aci
24	4244	61.2	1339	4 AAM78421	Human pro
25	4239.5	61.2	1306	4 AAB68218	Amino aci

26	4195.5	60.5	1297	2 AAY55932	Human ZC2
27	4021.5	58.0	1239	2 AAY55931	Human ZC1
28	3977	57.4	1212	4 AAE04368	Human kin
29	3977	57.4	1212	7 ADE25753	Human pro
30	3960.5	57.2	1165	3 AAB43016	Human ORF
31	3951.5	57.0	1233	2 AAY55954	Mouse STE
32	3872.5	55.9	1165	4 ABG17470	Novel hum
33	3872.5	55.9	1165	5 AAO18509	Human ins
34	3872.5	55.9	1165	6 ASU03499	Angiogene
35	3824	55.2	1175	4 ABG17468	Novel hum
36	3824	55.2	1175	5 AAU76512	Human HPK
37	3821	55.1	792	7 ADE34152	Human mis
38	3807.5	55.0	1655	4 ABG17466	Novel hum
39	3777.5	54.5	1135	3 AAY68784	Amino aci
40	3694.5	53.3	1219	4 ABG17469	Novel hum
41	2708.5	39.1	1109	2 AAY55953	Nematode
42	2280	32.9	510	3 AAB54144	Human pan
43	1862.5	26.9	1227	2 AAY55965	Full leng
44	1819.5	26.3	1581	5 AAE24145	Human kin
45	1686	24.3	425	4 AAB92792	Human pro

ALIGNMENTS

RESULT 1
ADE34150
ID ADE34150 standard; protein; 1312 AA.
XX AC ADE34150;
XX AC ADE34150;
DT 29-JAN-2004 (first entry)
XX DE Human misshapen/NIKS-related kinase, Mink3a.
XX KW Human; misshapen/NIKS-related kinase; Mink3a; enzyme; antiinflammatory;
KW immunosuppressive; cytostatic; germinal centre kinase;
KW C-JUN N-terminal kinase; JNK; extracellular signal response kinase; ERK;
KW growth factor induced-ERK activation; proliferation;
KW cell proliferation disorder; cell survival;
KW intracellular signal transduction; apoptosis; morphological change;
KW cell migration; gene therapy; inflammatory disease; autoimmune disease;
KW immunodeficiency; cancer.
XX OS Homo sapiens.
XX PN US2003077597-A1.
XX PD 24-APR-2003.
XX PF 19-OCT-2001; 2001US-00029115.
XX PR 19-OCT-2001; 2001US-00029115.
XX PA (LUOY/) LUO Y.
XX PA (FUCA/) FU C A.
XX PA (SHEN/) SHEN M.
XX PI Luo Y, Fu CA, Shen M;
XX DR WPI; 2003-635076/60.
XX DR N-PSDB; ADE34151.
XX PT New misshapen/NIKS-related kinase nucleic acids and proteins useful in
XX PT gene therapy and for treating disorders, e.g. acute and chronic
XX PT inflammatory diseases.
XX PS Claim 5; SEQ ID NO 2; 53pp; English.
XX CC The invention relates to a recombinant nucleic acid capable of
XX CC hybridising to a Human DNA encoding misshapen/NIKS-related kinase
XX CC (Mink3a, 3b and 3c, germinal centre kinase proteins) appearing as
XX CC ADE34151, ADE34153 and ADE34155, or at least 90% identity to them, or

CC their complements. Also included are a recombinant polypeptide at least
CC 95 % sequence identity to Mink3a, 3b or 3c (appearing as ADE34150,
CC ADE34152 and ADE34154), screening for a candidate bioactive agent capable
CC of modulating c-JUN N-terminal kinase (JNK) or extracellular signal
CC response kinase (ERK) phosphorylation or activity, screening for a
CC candidate bioactive agent capable of modulating growth factor induced-ERK
CC activation in a mammalian cell, screening for a candidate bioactive agent
CC capable of modulating proliferation in a mammalian cell, diagnosing a
CC mammalian cell proliferation disorder, a medicament for treating a
CC mammalian cell proliferation disorder and screening for a candidate agent
CC capable of modulating cell survival. The MINK3 (mishapen/NIRKs-related
CC kinase) nucleic acids are useful in the modulation of intracellular
CC signal transduction, cell proliferation, apoptosis, morphological change
CC and migration of mammalian cells. MINK3 nucleic acids and proteins are
CC specifically useful in gene therapy, and for treating, preventing or
CC diagnosing acute and chronic inflammatory diseases, autoimmune diseases
CC and diseases characterised by immunodeficiency. The compositions may also
CC be used to treat MINK3 dysfunction related disorders, e.g. cancer. The
CC nucleotide sequences may also be used as hybridisation probes, in
CC chromosome and gene mapping, and in generating antisense RNA and DNA. The
CC present sequence represents Mink3a.
XX
SQ Sequence 1312 AA;

Query Match 100.0%; Score 6929; DB 7; Length 1312;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGQYKGRHVKTGQLAAIKVMDYTE 60
DB 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGQYKGRHVKTGQLAAIKVMDYTE 60

QY 61 DEEEIKQBINLKKYSHRNATYTGAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120
DB 61 DEEEIKQBINLKKYSHRNATYTGAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120

QY 121 KGNALKEDCIAYICREILGLAHLHAHKVHHRDIKQNVLLTENAELVDFGVAQLDR 180
DB 121 KGNALKEDCIAYICREILGLAHLHAHKVHHRDIKQNVLLTENAELVDFGVAQLDR 180

QY 181 TVGRNTFTGTWYMAPEVIACDENPDATYDYSIDWSLIGITAIEMAGAPPLCDMHPMR 240
DB 181 TVGRNTFTGTWYMAPEVIACDENPDATYDYSIDWSLIGITAIEMAGAPPLCDMHPMR 240

QY 241 ALFLIPRNPPLKSKKSKKFIIDTCLITKYLSPPTTEOLLKFPFIIDOPTEROVRI 300
DB 241 ALFLIPRNPPLKSKKSKKFIIDTCLITKYLSPPTTEOLLKFPFIIDOPTEROVRI 300

QY 301 QLKDHIDRSRKRKGEKEETEYISGSEEDDSHGEGEPSSIMNVPGESTLRREFLRQQ 360
DB 301 QLKDHIDRSRKRKGEKEETEYISGSEEDDSHGEGEPSSIMNVPGESTLRREFLRQQ 360

QY 361 ENKNSSEALKQOQ 420
DB 361 ENKNSSEALKQOQ 420

QY 421 LOEKQOQ 480
DB 421 LOEKQOQ 480

QY 481 QOQ 540
DB 481 QOQ 540

QY 541 KPGSTGPEPPIQASGPGPGPLSQTPPMQRPVEPQEGPHKSLQDQPTRNLAAPASHDPD 600
DB 541 KPGSTGPEPPIQASGPGPGPLSQTPPMQRPVEPQEGPHKSLQDQPTRNLAAPASHDPD 600

QY 601 PAIPAPTATPSARGAVIRQNSDPTSEGGPGSPNPAPWRPDNEAPPKVPQRTSSATATLN 660
DB 601 PAIPAPTATPSARGAVIRQNSDPTSEGGPGSPNPAPWRPDNEAPPKVPQRTSSATATLN 660

QY 661 TSGAGGSRPAQAVRAPRPSNSAWQIYLQRRAGERTKPPGPPAQPQPPNASSNPDLRRS 720

DB 661 TSGAGGSRPAQAVRAPRPSNSAWQIYLQRRAGERTKPPGPPAQPQPPNASSNPDLRRS 720
QY 721 DPGWERSDSVLPASHGHLPOAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPGRPA 780
DB 721 DPGWERSDSVLPASHGHLPOAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPGRPA 780
QY 781 DFLVLLKERTLEAPRPPKAMDYSSSSSEVESSDEDEEGGPAEGSRDTPGGRSDGDT 840
DB 781 DFLVLLKERTLEAPRPPKAMDYSSSSSEVESSDEDEEGGPAEGSRDTPGGRSDGDT 840
QY 841 DSVSTMVVHDVEEITGTPPYGGGTVMVYQRTPEERNLLHADSNGYTNLPDVVQPSHSPT 900
DB 841 DSVSTMVVHDVEEITGTPPYGGGTVMVYQRTPEERNLLHADSNGYTNLPDVVQPSHSPT 900
QY 901 ENSKGQSPPSKDGSGDQSRGLVKAPGKSTFTMFDLGIYOPGSGSDSIPITALVGGSGT 960
DB 901 ENSKGQSPPSKDGSGDQSRGLVKAPGKSTFTMFDLGIYOPGSGSDSIPITALVGGSGT 960
QY 961 RLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKCRFNSILCAALGVNLLVGTENGL 1020
DB 961 RLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKCRFNSILCAALGVNLLVGTENGL 1020
QY 1021 MLLDRSGGKYGLIGRRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPE 1080
DB 1021 MLLDRSGGKYGLIGRRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPE 1080
QY 1081 VEKKQGTWTVGDMGCGGHYRVKRYKFLVIALKSSVEVYAWAPKPYHKFMAPKSFADL 1140
DB 1081 VEKKQGTWTVGDMGCGGHYRVKRYKFLVIALKSSVEVYAWAPKPYHKFMAPKSFADL 1140
QY 1141 PHRELLVDLTVEEGORLKVYSSAGFHAVDVDSGNSYDIYIPVHIOQITPHAIIFLPN 1200
DB 1141 PHRELLVDLTVEEGORLKVYSSAGFHAVDVDSGNSYDIYIPVHIOQITPHAIIFLPN 1200
QY 1201 TDGMELLCYDEDEGVYNTYGRRIIKDVVLOWGEMPTSVAYICSNQIMGWGEKAIEIRSV 1260
DB 1201 TDGMELLCYDEDEGVYNTYGRRIIKDVVLOWGEMPTSVAYICSNQIMGWGEKAIEIRSV 1260
QY 1261 TGHLDGVFMHKAQRLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
DB 1261 TGHLDGVFMHKAQRLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312

RESULT 2
AAO18508
ID AAO18508 standard; protein; 1295 AA.
XX
AC AAO18508;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human insulin receptor signaling modifier SEQ ID NO: 38.
XX
KW Human; insulin receptor signaling; insulin receptor signaling modifier;
KW ISM; diabetes; metabolic syndrome; antidiabetic.
XX
OS Homo sapiens.
XX
FN WO200255664-A2.
XX
PD 18-JUL-2002.
XX
PF 11-JAN-2002; 2002WO-US001048.
XX
PR 12-JAN-2001; 2001US-0261226P.
PR 12-JAN-2001; 2001US-0261303P.
PR 12-JAN-2001; 2001US-0261304P.
PR 12-JAN-2001; 2001US-0261335P.
PR 12-JAN-2001; 2001US-0261336P.
PR 12-JAN-2001; 2001US-0261361P.
PR 12-JAN-2001; 2001US-0261456P.
PR 12-JAN-2001; 2001US-0261457P.

PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261694P.
PR 12-JAN-2001; 2001US-0261695P.
PR 12-JAN-2001; 2001US-0261697P.
XX (EXEL-) EXELIXIS INC.
XX Seidel-Dugan C, Ferguson KC, Kidd T;
XX N-PSDB; AAL48627.
XX WPI; 2002-599664/64.
XX DR N-PSDB; AAL48627.
XX PT Identifying an insulin receptor signaling modulator, useful as drug
XX targets for treating diabetes or metabolic disorders, comprises
XX contacting an assay system comprising insulin receptor signaling
XX modifiers with a test agent.
XX Disclosure; Page 120-125; 232pp; English.
XX The present invention relates to a method of identifying a candidate
XX insulin receptor (INR) signaling modulating agent, involving contacting
XX an assay system comprising an insulin receptor signaling modifier (ISM)
XX polypeptide or nucleic acid with a test agent, and detecting a test agent
XX biased activity of the assay system. The method is useful for
XX identifying candidate INR signaling modulating agents. ISM genes may be
XX used as drug targets for treatment of disorders related to INR signaling
XX such as diabetes or metabolic syndrome. ISM nucleic acids and
XX polypeptides are useful for identifying and testing agents that modulate
XX ISM function and for other applications related to the involvement of ISM
XX in INR signaling, and for identifying subjects having a predisposition to
XX such diseases associated with INR signaling. The present sequence is an
XX ISM protein described in the exemplification of the invention
XX SQ Sequence 1295 AA;
Query Match 96.1%; Score 6660.5; DB 5; Length 1295;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;
QY 1 MGDPAKSLDDIDLSALRDPAGIFELVEVWNGTYGVYKGRHVKTQLAAIKVMDYTE 60
DB 1 MGDPAKSLDDIDLSALRDPAGIFELVEVWNGTYGVYKGRHVKTQLAAIKVMDYTE 60
QY 61 DEEBEIKQEIINMLKKYSHHRNIATYGAFFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
DB 61 DEEBEIKQEIINMLKKYSHHRNIATYGAFFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
QY 121 KGNALKEDCIAVICREILRLGLAHLAHKVIHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
DB 121 KGNALKEDCIAVICREILRLGLAHLAHKVIHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
QY 181 TVGRENFTGTYPWMAPEVACDENPDATYDRSDIWSIGITAIEMASGAPPLCDMHPR 240
DB 181 TVGRENFTGTYPWMAPEVACDENPDATYDRSDIWSIGITAIEMASGAPPLCDMHPR 240
QY 241 ALFLIPRNPPLRLSKSKWKKFIDFIDCLIKTYLSRPPTQQLKFFPIRQDPTQVR 300
DB 241 ALFLIPRNPPLRLSKSKWKKFIDFIDCLIKTYLSRPPTQQLKFFPIRQDPTQVR 300
QY 301 QUKOHIDSRKRGKEETEYYSGEEDDGHGEGEPSSIMNVPGESTLRREFLRQQ 360
DB 301 QUKOHIDSRKRGKEETEYYSGEEDDGHGEGEPSSIMNVPGESTLRREFLRQQ 360
QY 361 ENKNSSEALKKQQQQLQQQQOORDPEAHIKHLLHQRORRLEEOKEERRRVEEQORRERQK 420
DB 361 ENKNSSEALKKQQQQLQQQQOORDPEAHIKHLLHQRORRLEEOKEERRRVEEQORRERQK 420

QY 421 LOEKEQORRLLEDMOALRRERROAEREOEYKRYKQLEBEQORSERLQROLOQEHAYLKSLO 480
DB 421 LOEKEQORRLLEDMOALRRERROAEREOEYKRYKQLEBEQORSERLQROLOQEHAYLKSLO 480
QY 481 QQQQQQQQLQKQQQQQLLPGDKRPLHYHGRGNWNPADKPAWAREVEERTMNNKQNSPLAKS 540
DB 481 QQQQQQQQLQKQQQQQLLPGDKRPLHYHGRGNWNPADKPAWAREVEERTMNNKQNSPLAKS 540
QY 541 KPGSTGPEPPIPOASPGPPGSLSTQTPPMQRPVPEQEPGHK----- 580
DB 541 KPGSTGPEPPIPOASPGPPGSLSTQTPPMQRPVPEQEPGHKSLVAHRVPLKPYAAPVPRSQ 600
QY 581 SLQDQPTNLAAPFASHDDPPAIPATPATPSARGAVIRONSDDPTSEGPGSPNPAWVRP 640
DB 601 SLQDQPTNLAAPFASHDDPPAIPATPATPSARGAVIRONSDDPTSEGPGSPNPAWVRP 660
QY 641 DNEAPPKVPQRTSSIALNTSGAGSRPAQAVRAPRPSNSAWQIYLQRRARERTPKPPG 700
DB 661 DNEAPPKVPQRTSSIALNTSGAGSRPAQAVRA----- 695
QY 701 PPAQPPGPPNASSNPDLRRSDPGWERSDVLPAASHGLHPQAGSLERNRVGASSKLDSSPV 760
DB 696 -----SNPDLRRSDPGWERSDVLPAASHGLHPQAGSLERNRVGSSKPDSSPV 743
QY 761 LSPGNKAKPDHRSRPGRPADFVLLKERTLDEAPPPKAMDYSSSESESEDEDEEG 820
DB 744 LSPGNKAKPDHRSRPGRPADFVLLKERTLDEAPPPKAMDYSSSESESEDEDEEG 803
QY 821 EGGPAESGRDTPGGRSDGDTDSVSTMVVDVEEITGTQPPYGGGTMMVQVQRTPEBERNLH 880
DB 804 EGGPAESGRDTPGGRSDGDTDSVSTMVVDVEEITGTQPPYGGGTMMVQVQRTPEBERNLH 863
QY 881 ADSNGYTNLPDVVQPSHSPSTENSKQSPSPKQSGDYQSRGLVKAPGKSSFTMFVLDGIY 940
DB 864 ADSNGYTNLPDVVQPSHSPSTENSKQSPSPKQSGDYQSRGLVKAPGKSSFTMFVLDGIY 923
QY 941 QPGSGSDSIPITLVGEGEGLDQLDQYDVRKGSVNNVNTTRAHSETPEIRKYKRRNS 1000
DB 924 QPGSGSDSIPITLVGEGEGLDQLDQYDVRKGSVNNVNTTRAHSETPEIRKYKRRNS 983
QY 1001 EILCAALGVNLLVGTENGLMLDRSGGKGVYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
DB 984 EILCAALGVNLLVGTENGLMLDRSGGKGVYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
QY 1061 KLRVYLSWLRNKILHNDPEVEKKQGTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSWLRNKILHNDPEVEKKQGTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKPMATPKSFADLPHRPLVLDLTVEEQRLKVIYSSAGHAFVDDVDSGNSYDI 1180
DB 1104 YAWAPKPYHKPMATPKSFADLPHRPLVLDLTVEEQRLKVIYSSAGHAFVDDVDSGNSYDI 1163
QY 1181 YIPVHIQSQTIPHAIFLPTNDGMELLCYEDEGVVNTYGRRIIKDVVLQWGEPMPTSVAY 1240
DB 1164 YIPVHIQSQTIPHAIFLPTNDGMELLCYEDEGVVNTYGRRIIKDVVLQWGEPMPTSVAY 1223
QY 1241 ICSNQINGWGEKALTEIRSVEVTHGLDGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1300
DB 1224 ICSNQINGWGEKALTEIRSVEVTHGLDGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1283
QY 1301 FMTLNRNCIMNW 1312
DB 1284 FMTLNRNCIMNW 1295
RESULT 3
AAAY55933
ID AAAY55933 standard; protein; 1326 AA.
XX
AC AAAY55933;
XX
DT 18-FEB-2000 (first entry)

QY 121 KGNALKEDCIAVIREILRGLAHLHAHVKVIHRDIKQNVLLTENAIEVKLVDFGVGAQLDR 180
DB 121 KGNALKEDCIAVIREILRGLAHLHAHVKVIHRDIKQNVLLTENAIEVKLVDFGVGAQLDR 180
QY 181 TVGRENFTIGTYPWMAPEVACDENPDATYDVRSDIWSLIGITAIEMAGAPPLCDMHPWR 240
DB 181 TVGRENFTIGTYPWMAPEVACDENPDATYDVRSDIWSLIGITAIEMAGAPPLCDMHPWR 240
QY 241 ALFLIPRNPPLRLSKSKWKKFIDFIDTCLIKTYLSRPPTTEQLLKFPFIRDOPTERQVRI 300
DB 241 ALFLIPRNPPLRLSKSKWKKFIDFIDTCLIKTYLSRPPTTEQLLKFPFIRDOPTERQVRI 300
QY 301 QLKOHIDRSRKRGSKETEYYSGBEDDSDHGEPEPSSIMNVPGESTTLRREFLRLOQ 360
DB 301 QLKOHIDRSRKRGSKETEYYSGBEDDSDHGEPEPSSIMNVPGESTTLRREFLRLOQ 360
QY 361 ENKSNSEALKKQOQQLQOQOORDPPEAHIKHLHORORRIIEQKEERRRVEEQORREREOK 420
DB 361 ENKSNSEALKKQOQQLQOQOORDPPEAHIKHLHORORRIIEQKEERRRVEEQORREREOK 420
QY 421 LOEKEQORRLEDMQALREERERQAREQO-----EY 451
DB 421 LOEKEQORRLEDMQALREERERQAREQOYIHRLEBEQORLETLQOQLLEQOALLLEY 480
QY 452 KRKQLEEQORSRLQRLQOQEHAYILKSLQOQOQOQOQLKQOQOQLLPGDRKPLHYHGRGM 511
DB 481 KRKQLEEQORSRLQRLQOQEHAYILKSLQOQOQOQOQLKQOQOQLLPGDRKPLHYHGRGM 540
QY 512 NPADKPANAREVEERTNMKQNSPLAKSKPGSTGPEPPIPOASPGPGPLSOTPPMORP 571
DB 541 NPADKPANAREVEERTNMKQNSPLAKSKPGSTGPEPPIPOASPGPGPLSOTPPMORP 600
QY 572 VEPQEGPHK-----SLQDQTRNLAAFPASHDDPDAIPATATPS 611
DB 601 VEPQEGPHKSLVAHVPLKPYAAPVPRSQSLQDQTRNLAAFPASHDDPDAIPATATPS 660
QY 612 ARGAVIRQNSDPTSGPGSPNPAPWVRPDNEAPPKVPQRTSSITATNTSGAGSRPAQ 671
DB 661 ARGAVIRQNSDPTSGPGSPNPAPWVRPDNEAPPKVPQRTSSITATNTSGAGSRPAQ 720
QY 672 AVRARPNSAWQIYLQRAERGTGPKPGPPAQPFGPPNASSNPDLRRSDPGWERSDVL 731
DB 721 AVRA-----SNPDLRRSDPGWERSDVL 743
QY 732 PASHGHLPOAGSLERNRVGASSKLDSSPVLSPGNKAKPDHRSRGRPA-----DFV 783
DB 744 PASHGHLPOAGSLERNRVGVSSKPPSSPVLSPGNKAKPDHRSRGRPASVKRAIGEDFV 803
QY 784 LKERTLDEAPPPKAMDYSSSEBESSEDEEGEGGPAEGSRDTPGGRSDGDTDSV 843
DB 804 LKERTLDEAPPPKAMDYSSSEBESSEDEEGEGGPAEGSRDTPGGRSDGDTDSV 863
QY 844 STMVVDHVEEITGTOPPGGTMVQVTPPEERNLLHADSNGYTNLPDVQVPSHSPTENS 903
DB 864 STMVVDHVEEITGTOPPGGTMVQVTPPEERNLLHADSNGYTNLPDVQVPSHSPTENS 923
QY 904 KQSPSPSKDGSQSRGLVKAPEGKSSFTMFVDLGIYQPGSGSDIPITALVGGSGTRLD 963
DB 924 KQSPSPSKDGSQSRGLVKAPEGKSSFTMFVDLGIYQPGSGSDIPITALVGGSGTRLD 983
QY 964 QLYQVDRKGSVVNVNPTNTRAHSETPEIRKYKRRFNSSEILCAALGVNLLVGTENGLMLL 1023
DB 984 QLYQVDRKGSVVNVNPTNTRAHSETPEIRKYKRRFNSSEILCAALGVNLLVGTENGLMLL 1043
QY 1024 DRSGGKQVGLIGRRRFOQMDVLEGNLLITISGRNKLVRVYLSWLRNKLHNDPEVEK 1083
DB 1044 DRSGGKQVGLIGRRRFOQMDVLEGNLLITISGRNKLVRVYLSWLRNKLHNDPEVEK 1103
QY 1084 KQWTTVGDMECGHYRVVKYERIKFLVLTALKSSVEVWAPKPYHKFWAFKSPADLPHR 1143
DB 1104 KQWTTVGDMECGHYRVVKYERIKFLVLTALKSSVEVWAPKPYHKFWAFKSPADLPHR 1163

QY 1144 PLLVDLTVESGORKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSQTTPHAIIFLPNTDGG 1203
DB 1164 PLLVDLTVESGORKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSQTTPHAIIFLPNTDGG 1223
QY 1204 MEMLLCYEDEGVYNTYTGRIIKOVVLQWGMPTSVAVIYCSNQIMGWGEKAIEIRSVEIGH 1263
DB 1224 MEMLLCYEDEGVYNTYTGRIIKOVVLQWGMPTSVAVIYCSNQIMGWGEKAIEIRSVEIGH 1283
QY 1264 LDGVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRNCIMNW 1312
DB 1284 LDGVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRNCIMNW 1332
RESULT 8
ADE34154
ID ADE34154 standard; protein; 1276 AA.
XX
AC ADE34154;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human misshapen/NIKs-related kinase, Mink3c.
XX
KW Human; misshapen/NIKs-related kinase; Mink3c; enzyme; antiinflammatory;
KW immunosuppressive; cytostatic; germinal centre kinase;
KW c-JUN N-terminal kinase; JNK; extracellular signal response kinase; ERK;
KW growth factor induced-ERK activation; proliferation;
KW cell proliferation disorder; cell survival;
KW intracellular signal transduction; apoptosis; morphological change;
KW cell migration; gene therapy; inflammatory disease; autoimmune disease;
KW immunodeficiency; cancer.
XX
OS Homo sapiens.
XX
PN US2003077597-A1.
XX
PD 24-APR-2003.
XX
PF 19-OCT-2001; 2001US-00029115.
XX
PR 19-OCT-2001; 2001US-00029115.
XX
PA (LUOY/) LUO Y.
PA (FUCA/) FU C A.
PA (SHEN/) SHEN M.
XX
PI Luo Y, Fu CA, Shen M;
XX
PI WPI; 2003-635076/60.
DR N-PSDB; ADE34155.
XX
PT New misshapen/NIKs-related kinase nucleic acids and proteins useful in
PT gene therapy and for treating disorders, e.g. acute and chronic
PT inflammatory diseases.
XX
PS Claim 5; SEQ ID NO 6; 53pp; English.
XX
CC The invention relates to a recombinant nucleic acid capable of
CC hybridising to a Human DNA encoding misshapen/NIKs-related kinase
CC (Mink3a, 3b and 3c, germinal centre kinase proteins) appearing as
CC ADE34151, ADE34153 and ADE34155, or at least 90% identity to them, or
CC their complements. Also included are a recombinant polypeptide at least
CC 95 % sequence identity to Mink3a, 3b or 3c (appearing as ADE34150,
CC ADE34152 and ADE34154), screening for a candidate bioactive agent capable
CC of modulating c-JUN N-terminal kinase (JNK) or extracellular signal
CC response kinase (ERK) phosphorylation or activity, screening for a
CC candidate bioactive agent capable of modulating growth factor induced-ERK
CC activation in a mammalian cell, screening for a candidate bioactive agent
CC capable of modulating proliferation in a mammalian cell, diagnosing a
CC mammalian cell proliferation disorder, a medicament for treating a
CC mammalian cell proliferation disorder and screening for a candidate agent
CC capable of modulating cell survival. The MINK3 (misshapen/NIKs-related
CC kinase) nucleic acids are useful in the modulation of intracellular

876 KNTLHADSNGIINLFDDVVFQSHSFIENTSRKGQSPFFSNKGSGDIQSRKGLVNAFGKSSFIMFV :333

PT polypeptide for diagnosing and treating pathological disorders, such as
PT Parkinson's disease and for use in pharmacogenomics.
XX Claim 1; Page 68; 189pp; English.
PS The invention relates to novel human polypeptides referred as NOV-X and
XX their corresponding nucleic acid sequences. NOV-X collectively include
CC NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-
CC 3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides and
CC NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin
CC inhibitor-like polypeptides. NOV-X is used to identify a potential
CC therapeutic agent that can modulate its activity and can be used for
CC treating a pathology related to aberrant expression or aberrant
CC physiological interactions of NOV-X. NOV-X or its DNA is used to
CC determine the presence or predisposition to a disease associated with
CC altered levels of NOV-X. NOV-X, its DNA and its antibody are used to
CC treat or prevent a pathology associated with NOV-X. The pathological
CC states that can be treated or prevented are haematopoietic, cancer,
CC immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC fertility disorders. NOV-X and its DNA are used in pharmacogenomics for
CC predictive medicine. NOV-X DNA is used in gene therapy. The present
CC sequence is human novel STE20-like protein, NOV-3d. NOV-3d gene is
XX located at chromosome 17
SQ Sequence 1244 AA;
Query Match 91.1%; Score 6311; DB 4; Length 1244;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 1218; Conservative 3; Mismatches 15; Indels 84; Gaps 4;
QY 1 MGDPAARSLLDDIDLSALRDPAGIFELVEVNGTYGYQYGRHVKTQGLAAIKVMDVTE 60
DB 1 MGDPAARSLLDDIDLSALRDPAGIFELVEVNGTYGYQYGRHVKTQGLAAIKVMDVTE 60
QY 61 DEEEIKOEINMLKKYSHRNATYTGAFIKSPGNDQDLWLMVEFCGAGSVTLVKNT 120
DB 61 DEEEIKOEINMLKKYSHRNATYTGAFIKSPGNDQDLWLMVEFCGAGSVTLVKNT 120
QY 121 KGNALKEDCIAYICREILRGLAHLHAHVHRIHDKQNVLLTENAELVDFGSAQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHVHRIHDKQNVLLTENAELVDFGSAQLDR 180
QY 181 TVGRNTFTGTPYMAPEVIACDENPDATYDYSIWSLIGTATIAEMAGAPPLCDMHPMR 240
DB 181 TVGRNTFTGTPYMAPEVIACDENPDATYDYSIWSLIGTATIAEMAGAPPLCDMHPMR 240
QY 241 ALFLIPRPPPLKSKKSKKFIIDTCLIKTYLSRPTTEQLLKFPPIRDOPTEROVRI 300
DB 241 ALFLIPRPPPLKSKKSKKFIIDTCLIKTYLSRPTTEQLLKFPPIRDOPTEROVRI 300
QY 301 QLKDHIDRSKRKRGEKEETEYSGSEEDSDSHGEGEPSSIMNVPGESTLRRREFLRQQ 360
DB 301 QLKDHIDRSKRKRGEKEETEYSGSEEDSDSHGEGEPSSIMNVPGESTLRRREFLRQQ 360
QY 361 ENKNSSEALKQOQQOQQOQQOQDPEAHIKHLHQRORRIIEQKEERRRVEEQORREQRK 420
DB 361 ENKNSSEALKQOQQOQQOQQOQDPEAHIKHLHQRORRIIEQKEERRRVEEQORREQRK 420
QY 421 LOEKEQORLEDMOALREBERERQAREGEYKRLQEQORSERLQRLQOEHAYLKSLO 480
DB 421 LOEKEQORLEDMOALREBERERQAREGEYKRLQEQORSERLQRLQOEHAYLKSLO 480
QY 481 OQOQQOQQOQQOQQOQQOQQOQQOQDPEAHIKHLHQRORRIIEQKEERRRVEEQORREQRK 540
DB 481 OQOQQOQQOQQOQQOQQOQQOQQOQDPEAHIKHLHQRORRIIEQKEERRRVEEQORREQRK 540
QY 541 KPGSTGPPPIPAQSPGPGPLSPQTPPMQRPVPEQEGPHKSLQDQPTRNLAAPASHDPD 600
DB 541 KPGSTGPPPIPAQSPGPGPLSPQTPPMQRPVPEQEGPHKSLQDQPTRNLAAPASHDPD 600
QY 528 -----VPLKPYAAPVPRS-----QSLQDQPTRNLAAPASHDPD 561
DB 528 -----VPLKPYAAPVPRS-----QSLQDQPTRNLAAPASHDPD 561
QY 601 PAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPAPWVRPDNEAPPKVPQORTSSATATLN 660
DB 601 PAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPAPWVRPDNEAPPKVPQORTSSATATLN 660

Db 562 PAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPAPWVRPDNEAPPKVPQORTSSATATLN 621
QY 661 TSGAGGSRPAQAVRAPRPSNSAWQIYLQRAERGTGPKPPGPPAQPAGPPNASSNPDLRRS 720
Db 622 TSGAGGSRPAQAVRA-----SNPDLRRS 644
QY 721 DPGHERSDSLPASHGHLPOAGSLERNRVGASSKLDSPLVSPGNKAKPDHRSRPGRPA 780
Db 645 DPGHERSDSLPASHGHLPOAGSLERNRVGSSKPDSPVLSPGNKAKPDHRSRPGRPA 704
QY 781 -----DFVLLKERTLDEAPRPPKAMDYSSSESEVESEDEDEEGEGGPAEGSRDTP 832
Db 705 SYKRAIGEDFVLLKERTLDEAPRPPKAMDYSSSESEVESEDEDEEGEGGPAEGSRDTP 764
QY 833 GGRSDGTDTSVSTVMVHDVEITGTQPPYGGTVMVQRTPEERNLLHADNSGYTNLPDV 892
Db 765 GGRSDGTDTSVSTVMVHDVEITGTQPPYGGTVMVQRTPEERNLLHADNSGYTNLPDV 824
QY 893 VOPSHSPTEKSGQSPKSGDGYQSRGLVKAPKSSFTMPVDLGIYQPGSGDSIPIT 952
Db 825 VOPSHSPTEKSGQSPKSGDGYQSRGLVKAPKSSFTMPVDLGIYQPGSGDSIPIT 884
QY 953 ALVGGEGTRLDQLQYDVRKGSVNVNPTNTTRAHSETPEIRKYKRFNSEILCAALWGVNL 1012
Db 885 ALVGGEGTRLDQLQYDVRKGSVNVNPTNTTRAHSETPEIRKYKRFNSEILCAALWGVNL 944
QY 1013 LVGTENGLMLDLSRGGQKGYGLIGRRRFQOQMDVLLEGLNLLITISGKRNLKRVYYLSWLRN 1072
Db 945 LVGTENGLMLDLSRGGQKGYGLIGRRRFQOQMDVLLEGLNLLITISGKRNLKRVYYLSWLRN 1004
QY 1073 KILHNDPEVEKKQWTTVGMGCGHYRVVKYERIKFLVIALKSSVEYVAPKPYHKFM 1132
Db 1005 KILHNDPEVEKKQWTTVGMGCGHYRVVKYERIKFLVIALKSSVEYVAPKPYHKFM 1064
QY 1133 AFKSFADLPHRPLVLDLTVBEGQRLKVIYSGSAGFAHVDVDSGNSYDIYIPVHQSOQITP 1192
Db 1065 AFKSFADLPHRPLVLDLTVBEGQRLKVIYSGSAGFAHVDVDSGNSYDIYIPVHQSOQITP 1124
QY 1193 HAIIFLPNTDGMELLCYDEGEVYVNTYGRIIKDVVLQWGMPTSVAYICSNQIMGWGEK 1252
Db 1125 HAIIFLPNTDGMELLCYDEGEVYVNTYGRIIKDVVLQWGMPTSVAYICSNQIMGWGEK 1184
QY 1253 AIEIRSVETGHLQVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRCINMW 1312
Db 1185 AIEIRSVETGHLQVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRCINMW 1244
RESULT 10
AAE10613
ID AAE10613 standard; protein; 1273 AA.
XX
AC AAE10613;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human novel STE20-like protein, NOV-3c.
XX
KW Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour;
KW trypsin inhibitor-like protein; gene therapy; haematopoietic; illness;
KW immunological disorder; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic;
KW human immunodeficiency virus; HIV; fertility disorder; neuroprotective;
KW cytosolic; anti-infective; cancer; chromosome 17;
XX NOV-3c protein.
OS Homo sapiens.
XX
FN WO200162928-A2.
XX
PD 30-AUG-2001.
XX
PF 26-FEB-2001; 2001WO-US0006151.
XX

XX Novel human diagnostic protein #24011.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX N-PSDB; AAS88207.
DR WPI; 2001-639362/73.
XX N-PSDB; AAS88207.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 54379; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ASG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1701 AA;
Query Match 81.6%; Score 5651; DB 4; Length 1701;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1112; Conservative 8; Mismatches 22; Indels 212; Gaps 9;
QY 1 MGDPAASLDDIDLSALRDPAGIFELVEVNGTYGQVYKGRHVKTGOLAAIKVMDYTE 60
DB 1 MGDPAASLDDIDLSALRDPAGIFELVEVNGTYGQVYKGRHVKTGOLAAIKVMDYTE 60
QY 61 DEEEIKOEINMLKYSHHRNATYTGAFIKKSPGNDQDLWLVNVEFCAGSVTLVKNT 120
DB 61 DEEEIKOEINMLKYSHHRNATYTGAFIKKSPGNDQDLWLVNVEFCAGSVTLVKNT 120
QY 121 KGNALKEDCIAYICREILRGLAHLHAHKVYHRDIKQGNVLLTENAEVKLVDFGVSAQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHKVYHRDIKQGNVLLTENAEVKLVDFGVSAQLDR 180
QY 181 TVGRNTFTGTPYMAPEVIACDENPDATYDRSDIWSLIGITAIEMAGAPPLCDHMPWR 240

DB 181 TVGRNTFTGTPYMAPEVIACDENPDATYDRSDIWSLIGITAIEMAGAPPLCDHMPWR 240
QY 241 ALFLIPNPPRLKSKWSKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRQOFTERQVRI 300
DB 241 ALFLIPNPPRLKSKWSKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRQOFTERQVRI 300
QY 301 QKDHIDRSKRKGEKETETEYEGSEEDDSHGEGEPSSIMNVPGSTLRRREFLRQQ 360
DB 301 QKDHIDRSKRKGEKETETEYEGSEEDDSHGEGEPSSIMNVPGSTLRRREFLRQQ 360
QY 361 ENKSNSEALKQOQOQLQOQOQDPEAHIKHLHQRRIEEQKEERRRVEEQRREREQK 420
DB 361 ENKSNSEALKQOQOQLQOQOQDPEAHIKHLHQRRIEEQKEERRRVEEQRREREQK 420
QY 421 LOEKEQQRRLDMQALRRREERRQAEREQ-----EY 451
DB 421 LOEKEQQRRLDMQALRRREERRQAEREQYIRHLEBEERQLETLQOQLLQEQALLLEY 480
QY 452 KRKOLEBOROSERLQORLOQEHAYLKSLOOQOQOQOQOQOQOQLLPGDRKPLVHYGRGM 511
DB 452 KRKOLEBOROSERLQORLOQEHAYLKSLOOQOQOQOQOQOQOQLLPGDRKPLVHYGRGM 511
QY 512 NPADKPAREVEREERTMKNKQONSPLAKSKPGSTGPEPPIPOASPGPPGSLSTPPMPORP 571
DB 512 NPADKPAREVEREERTMKNKQONSPLAKSKPGSTGPEPPIPOASPGPPGSLSTPPMPORP 571
QY 541 NPADKPAREVEREERTMKNKQONSPLAKSKPGSTGPEPPIPOASPGPPGSLSTPPMPORP 600
DB 541 NPADKPAREVEREERTMKNKQONSPLAKSKPGSTGPEPPIPOASPGPPGSLSTPPMPORP 600
QY 572 VEPOEGPHK-----SILOQDPTNLAAPFASHDPPDPAIPAPTATPS 611
DB 572 VEPOEGPHKSLVAHRVPLKPYAAPVPRSQSLQDPTNLAAPFASHDPPDPAIPAPTATPS 660
QY 612 ARGAVIQNSDPTSEGPGSPNPAPVRPONEAPKVPQORTSSATATANTSGAGSRAQ 671
DB 612 ARGAVIQNSDPTSEGPGSPNPAPVRPONEAPKVPQORTSSATATANTSGAGSRAQ 671
QY 661 ARGAVIQNSDPTSEGPGSPNPAPVRPONEAPKVPQORTSSATATANTSGAGSRAQ 720
DB 661 ARGAVIQNSDPTSEGPGSPNPAPVRPONEAPKVPQORTSSATATANTSGAGSRAQ 720
QY 672 AVRAPRNSAWQIYLQRRARERTPKPPGPPAQPQPPNPNASNDLRRSDPGWERSDSVL 731
DB 672 AVRA-----SNPDLRRSDPGWERSDSVL 743
QY 732 PASHGLHLPQAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPRRA-----DFV 783
DB 732 PASHGLHLPQAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPRRA-----DFV 783
QY 744 PASHGLHLPQAGSLERNRVGSSKPDSSPVLSPGNKAKPDDHRSRPRRA-----DFV 803
DB 744 PASHGLHLPQAGSLERNRVGSSKPDSSPVLSPGNKAKPDDHRSRPRRA-----DFV 803
QY 784 LLKERTLDEAPRPKKAWDYSSSSSEVESSDEDEEGGPAEGSRDTPGSGSDGTDUSV 843
DB 784 LLKERTLDEAPRPKKAWDYSSSSSEVESSDEDEEGGPAEGSRDTPGSGSDGTDUSV 843
QY 804 LLKERTLDEAPRPKKAWDYSSSSSEVESSDEDEEGGPAEGSRDTPGSGSDGTDUSV 863
DB 804 LLKERTLDEAPRPKKAWDYSSSSSEVESSDEDEEGGPAEGSRDTPGSGSDGTDUSV 863
QY 844 STMVVDVEEITGTPPYGGGTVMVQRTPEERNLLHADSNGYTNLPDVQPSHSPTENS 903
DB 844 STMVVDVEEITGTPPYGGGTVMVQRTPEERNLLHADSNGYTNLPDVQPSHSPTENS 903
QY 864 STMVVDVEEITGTPPYGGGTVMVQRTPEERNLLHADSNGYTNLPDVQPSHSPTENS 923
DB 864 STMVVDVEEITGTPPYGGGTVMVQRTPEERNLLHADSNGYTNLPDVQPSHSPTENS 923
QY 904 KGQSPPSKDGSDYQSRGLVKAPCKSSFTMFVDLG1YQPGSGDSIPITALVGGEGTRLD 963
DB 904 KGQSPPSKDGSDYQSRGLVKAPCKSSFTMFVDLG1YQPGSGDSIPITALVGGEGTRLD 963
QY 924 KGQSPPSKDGSDYQSRGLVKAPCKSSFTMFVDLG1YQPGSGDSIPITALVGGEGTRLD 983
DB 924 KGQSPPSKDGSDYQSRGLVKAPCKSSFTMFVDLG1YQPGSGDSIPITALVGGEGTRLD 983
QY 964 QLOYDVRKGSVNVNPTNTRAHSETPEIRKVKPFNSEILCAALWGNLLVGTENGLMLL 1023
DB 964 QLOYDVRKGSVNVNPTNTRAHSETPEIRKVKPFNSEILCAALWGNLLVGTENGLMLL 1023
QY 984 QLOYDVRKGSVNVNPTNTRAHSETPEIRKVKPFNSEILCAALWGNLLVGTENGLMLL 1043
DB 984 QLOYDVRKGSVNVNPTNTRAHSETPEIRKVKPFNSEILCAALWGNLLVGTENGLMLL 1043
QY 1024 DRSGQGVYGLIGRRRFFQOMDVLGLNLLITISGRNKLRYVYLSWLNKILHNDPVEK 1083
DB 1024 DRSGQGVYGLIGRRRFFQOMDVLGLNLLITISGRNKLRYVYLSWLNKILHNDPVEK 1083
QY 1044 GPKWAG-----GKNKLRYVYLSWLNKILHNDPVEK 1077
DB 1044 GPKWAG-----GKNKLRYVYLSWLNKILHNDPVEK 1077
QY 1084 KQGWTTVDMEGCGHYRVVYKRIKFLVIALKSSVEYVAMAPKPYHKFMFKSFADLPHR 1143
DB 1084 KQGWTTVDMEGCGHYRVVYKRIKFLVIALKSSVEYVAMAPKPYHKFMFKSFADLPHR 1143
QY 1078 KQGWTTVDMEGCGHYRVVYKRIKFLVIALKSSVEY----- 1114
DB 1078 KQGWTTVDMEGCGHYRVVYKRIKFLVIALKSSVEY----- 1114
QY 1144 PLLVDLTVBEGQRLKVIYGSAGHFAVDVDSGNSYDIYIPVHQSIQTPHAIIFLPNTDG 1203
DB 1144 PLLVDLTVBEGQRLKVIYGSAGHFAVDVDSGNSYDIYIPVHQSIQTPHAIIFLPNTDG 1203
QY 1115 -----ARLPHAIIFLPNTDG 1130
DB 1115 -----ARLPHAIIFLPNTDG 1130
QY 1204 MEMLLCYDEGVVYNTYGRYIKDVLVQWGEEMPTSVAYICSNQIMGWGKAIEIRSVETGH 1263
DB 1204 MEMLLCYDEGVVYNTYGRYIKDVLVQWGEEMPTSVAYICSNQIMGWGKAIEIRSVETGH 1263
QY 1131 -----RRTKVTVV--NTLMGWGKAIEIRSVETGH 1165
DB 1131 -----RRTKVTVV--NTLMGWGKAIEIRSVETGH 1165

```

QY 1264 LDGVFMHKAQRLKFLCERNDKVFFASVSGSS 1297
DB 1166 LDGVFMHKAQRLKFLCERNDKVFFASVSGSS 1199

RESULT 12
ID AAB68221
XX AAB68221 standard; protein; 1324 AA.
AC AAB68221;
XX 09-JUL-2001 (first entry)
XX Amino acid sequence of cell cycle protein Thik isoform 6.
DE Cell cycle protein; Thik; germinal center kinase; Nck; Traf2; cell cycle;
KW tumour necrosis factor receptor associated factor 2; gene therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 1324
FT /note= "Glx encoded by TAA"
XX WO200129197-A2.
XX 26-APR-2001.
XX 23-OCT-2000; 2000WO-US041455.
XX 21-OCT-1999; 99US-00425324.
XX (RIGS-) RIGEL PHARM INC.
XX Luo Y, Fu CA, Shen M;
XX WPI; 2001-300333/31.
XX N-PSDB; AAF84944.
XX Novel germinal center kinase cell cycle polypeptide useful for screening
XX modulators of cell cycle and for use in diagnostics and therapeutics.
XX Claim 14; Fig 33; 96pp; English.
XX The present sequence represents an isoform of a human cell cycle protein
XX designated Thik. Thik is a member of the germinal center kinase family,
XX Thik binds to tumour necrosis factor (TNF) receptor associated factor 2
XX (Traf2) or the adapter protein Nck. Thik polypeptides and polynucleotides
XX are useful for screening for compounds that are modulators of cell cycle,
XX and in therapeutics. The Thik polynucleotide is also useful in gene
XX therapy
XX Sequence 1324 AA;

Query Match 63.1%; Score 4370.5; DB 4; Length 1324;
Best Local Similarity 64.7%; Pred. No. 9.7e-252;
Matches 899; Conservative 137; Mismatches 210; Indels 143; Gaps 31;

QY 1 MGDPAAPASLDDIDLSALRDPAGIFELVEVNGNGYGVYKGRHVKTGQLAAIKMVDYTG 60
DB 1 MASDSPASLDDIDLSALRDPAGIFELVEVNGNGYGVYKGRHVKTGQLAAIKMVDYTG 60
QY 61 DEEEIKOBINMLKKYSHRNATATYGAFFIKKPPGNDQDLWLVMFPGAGSVTDLVKNT 120
DB 61 DEEEIKOBINMLKKYSHRNATATYGAFFIKKPPGNDQDLWLVMFPGAGSVTDLVKNT 120
QY 121 KGNALKECDIAYICREILRGLAHLHAHKVYIHRDIKQNVLLTENAEVKLVDFGVSAQLDR 180
DB 121 KGNLKEEWIAYICREILRGLSHLHQHKVYIHRDIKQNVLLTENAEVKLVDFGVSAQLDR 180
QY 181 TVGRNTFTGTPYMAPEVIACDENPDATYDFKSLMSLGITAIEMAGAPPLCDMHPMR 240
DB 181 TVGRNTFTGTPYMAPEVIACDENPDATYDFKSLMSLGITAIEMAGAPPLCDMHPMR 240

```


QY 1095 GCCHYRVVYKRIKELVIALKSSVEYAWAPKYHFKFAKSPADLPHPRLVLDLTVBEG 1154
 DB 1135 GCCHYRVVYKRIKELVIALKNAVIYAWAPKYHFKFAKSPADLQHKPLVLDLTVBEG 1194
 QY 1155 QRLKVIYSGSAGFAVDVDSGNSYDIYIPVHQSOITTHAIIIFLPTDGMMLLCYDEBG 1214
 DB 1195 QRLKVIYSGSHTGFHVIDVDSGNSYDIYIPSHIQGNITPHAVILPKTDGMMLLCYDEBG 1254
 QY 1215 VYVNTYGRIRIKDVLQWGMPTSVAYICSNQIMGKGAIEIRSVETGHLGDFVHKRAQ 1274
 DB 1255 VYVNTYGRITKDWLQWGMPTSVAYIHSNQIMGKGAIEIRSVETGHLGDFVHKRAQ 1314
 QY 1275 RLKFLCERNDKVFFASVRSVSGSSOVFFMTLRNCITMNV 1312
 DB 1315 RLKFLCERNDKVFFASVRSVSGSSOVFFMTLRNCSMNV 1352

RESULT 15

AAV85263

ID AAV85263 standard; protein; 1360 AA.

XX

AC AAV85263;

XX 29-JUN-2000 (first entry)

XX

DE Human protein kinase KIAA0551 amino acid sequence.

XX

KW Protein kinase; KIAA0551; neuropathy; neuropathic pain; inflammation;

XX

KW chronic pain; neurodegenerative disorder; neurotraumatic disorder;

XX

KW Parkinson's disease; Alzheimer's disease; ischaemic disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT

FT Misc-difference 290

XX

XX WO200015805-A1.

XX

XX 23-MAR-2000.

XX

PF 10-SEP-1999; 99WO-GB003017.

XX

PR 10-SEP-1999; 98GB-00019779.

XX

PR 29-MAR-1999; 98GB-00007261.

XX

XX (SMTK) SMITHKLINE BEECHAM PLC.

XX

XX Bingham S, Case P, Lawson SN, Newton RA, Rausch OL, Reith AD;

PI

PI Sanger GJ;

XX

XX WPI; 2000-271443/23.

XX

XX N-PSDB; AAA10669.

XX

XX Isolated human KIAA0551 polynucleotide and polypeptide, useful for

PT

PT treating e.g. neuropathies, neuropathic pain, inflammatory and chronic

PT

PT pain and neurodegenerative conditions.

XX

XX Claim 2; Page 41; 48pp; English.

XX

XX This sequence represents the human protein kinase KIAA0551 amino acid

CC

CC sequence. The nucleotide sequence was isolated from a human foetal brain

CC

CC cDNA library. The rat KIAA0551 mRNA is upregulated in dorsal root ganglia

CC

CC during sciatic neuropathy (a procedure accompanied by increased

SQ Sequence 1360 AA;

Query Match 62.6%; Score 4336; DB 3; Length 1360;

Best Local Similarity 63.0%; Pred. No. 1.1e-249;

Matches 898; Conservative 138; Mismatches 210; Indels 180; Gaps 32;

QY 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNTYGOVYKGRHVKTGQLAAIKVMDVTE 60

DB 1 MASDPAARSDEIDLSALRDPAGIFELVEVVGNTYGOVYKGRHVKTGQLAAIKVMDVTG 60

QY 61 DEBEIKQEIINMLKYSHRNIATYYGAFIKKSPGNDQDLWMEFCGAGSVTDLVKNT 120

DB 61 DEBEIKQEIINMLKYSHRNIATYYGAFIKKSPGNDQDLWMEFCGAGSVTDLVKNT 120

QY 121 KGNALKECTIAYICREILRGLAHVHAKVHRDIKGNVLLTENAELVKLVDFGVSQAQLDR 180

DB 121 KGNALKECTIAYICREILRGLAHVHAKVHRDIKGNVLLTENAELVKLVDFGVSQAQLDR 180

QY 181 TVGRNTFIGTPYMAPEVIACDENPDATYDYSIDISLGITAJEMAGAPLDCMHMPMR 240

DB 181 TVGRNTFIGTPYMAPEVIACDENPDATYDYSIDISLGITAJEMAGAPLDCMHMPMR 240

QY 241 ALFLIPRNPRLSKKWSKKFIDFIDTCLIKTYLSRPPTRQLLKFPPIRQPTERQVRI 300

DB 241 ALFLIPRNPRLSKKWSKKFIDFIDTCLIKTYLSRPPTRQLLKFPPIRQPTERQVRI 300

QY 301 QLKDHIDRSRKRGEETEYEGSGEEDDSHGEPESSIMNVPGESTLRRRFLRLQQ 360

DB 301 QLKDHIDRSRKRGEETEYEGSGEEDDSHGEPESSIMNVPGESTLRRRFLRLQQ 360

QY 361 ENKENSEALKQOQLOQOQOQDPEAHIKHLLHQRRIEKEEKEERRRVEEQOREREOR 420

DB 361 ENKENSEALKQOQLOQOQOQDPEAHIKHLLHQRRIEKEEKEERRRVEEQOREREOR 420

QY 421 LOEKEQORRLDQALRREERROAREQ-----EY 451

DB 421 LOEKEQORRLDQALRREERROAREQ-----EY 451

QY 452 KRKOLEEOROSERLQRLQOEHAYKLSLQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 511

DB 452 KRKOLEEOROSERLQRLQOEHAYKLSLQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 511

QY 476 KRKOLEEORQAERLQRLQOERDYLVSLOHQORQ-----PVEKKPLVHYKEGM 525

DB 476 KRKOLEEORQAERLQRLQOERDYLVSLOHQORQ-----PVEKKPLVHYKEGM 525

QY 512 NPADKPAAWEVEERTMKNQONSPLAKSKGSGTGPPIPAQSPG-----PPCPLSQT 565

DB 512 NPADKPAAWEVEERTMKNQONSPLAKSKGSGTGPPIPAQSPG-----PPCPLSQT 565

QY 526 SPSEKPAWAKEVEERSRLNR-QSSPAMPHKVNKRISDFNLPFRSEFSISGVQP---ART 581

DB 526 SPSEKPAWAKEVEERSRLNR-QSSPAMPHKVNKRISDFNLPFRSEFSISGVQP---ART 581

QY 566 PPMORPVEPO-----EGP-----HKSLOQDQPTRNLAAFPASHDPPAIPAPTATPSA 612

DB 566 PPMORPVEPO-----EGP-----HKSLOQDQPTRNLAAFPASHDPPAIPAPTATPSA 612

QY 582 PPMORPVEPO-----EGP-----HKSLOQDQPTRNLAAFPASHDPPAIPAPTATPSA 612

DB 582 PPMORPVEPO-----EGP-----HKSLOQDQPTRNLAAFPASHDPPAIPAPTATPSA 612

QY 613 RGAVIRQNSDPTSEGPSP-----NPPAWVRPDNEAPPKVPQRTSSITALTNTSGAGG 666

DB 613 RGAVIRQNSDPTSEGPSP-----NPPAWVRPDNEAPPKVPQRTSSITALTNTSGAGG 666

QY 632 RVEMPRQNSDPTSENP-PLPTRIEKFDRSSWLQGEEDIPPKVPQRTTSSIPALARKNSPG 690

DB 632 RVEMPRQNSDPTSENP-PLPTRIEKFDRSSWLQGEEDIPPKVPQRTTSSIPALARKNSPG 690

QY 667 SRPAQVAPRPRNSAMQIVLORRAERGTPKPPPPAQPFGPPNASSNPDLRRSDPGWE- 725

DB 667 SRPAQVAPRPRNSAMQIVLORRAERGTPKPPPPAQPFGPPNASSNPDLRRSDPGWE- 725

QY 691 NGSALG---PLGSG-----QPTRASNPDLRRTEPILES 720

DB 691 NGSALG---PLGSG-----QPTRASNPDLRRTEPILES 720

QY 726 -----RSDSVLPASH-----GHLP---QAGSLERNVGCASKLSSPVL---SPGNKA 767

DB 726 -----RSDSVLPASH-----GHLP---QAGSLERNVGCASKLSSPVL---SPGNKA 767

QY 721 PLQRTSSGSSSSSTPSSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSG 779

DB 721 PLQRTSSGSSSSSTPSSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQSGSQ 779

QY 768 KPDDHR---SRGPRPADFVL-----LKERTLDEAPRPPPKAMDYSSSSSEEVERS 813

DB 768 KPDDHR---SRGPRPADFVL-----LKERTLDEAPRPPPKAMDYSSSSSEEVERS 813

QY 780 KPESRDITRPSRASYKKAIDEDLTALAKELRELREIETNRPMKVTYDSSSESESS 839

DB 780 KPESRDITRPSRASYKKAIDEDLTALAKELRELREIETNRPMKVTYDSSSESESS 839

QY 814 EDDDEEGEGGPAEGS---RDTG-----GGRSDGDTDSVSTWVVDVDEITGTQPPYGG-- 863

DB 814 EDDDEEGEGGPAEGS---RDTG-----GGRSDGDTDSVSTWVVDVDEITGTQPPYGG-- 863

QY 840 EEEDEGESETHDGTVAVSDIPRLIPTGAPSGNEQYVNVGMVGTGHLG--TSHADSFGSGI 897

DB 840 EEEDEGESETHDGTVAVSDIPRLIPTGAPSGNEQYVNVGMVGTGHLG--TSHADSFGSGI 897

QY 864 ---GTMVQVQTPPEERNLLHADNSNGYT---NLPDVQVPSHS---PTENSKGQSPPSKD- 912

DB 864 ---GTMVQVQTPPEERNLLHADNSNGYT---NLPDVQVPSHS---PTENSKGQSPPSKD- 912

QY 898 SREGTLMIRETSGEKGSGHSDNGFAGHINLPDLVQOQSHSPAGTPTIEGLGRVSTHSEQM 957

DB 898 SREGTLMIRETSGEKGSGHSDNGFAGHINLPDLVQOQSHSPAGTPTIEGLGRVSTHSEQM 957

Qy	913	GS	GDYQSRGLVAPGKSGSFTMEVDLGIYQPGS-----GDSIPITALVGGEGTRLDQLOY	967
Db	958	DS	GTGKG---SSTKASFTFVDPVQVTSPTDEDEDESSAAALFTSELLROEQAKL	1014
Qy	968	-DVR	KSVVNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRS	1026
Db	1015	NEAR	KISVVNVNPTNIRPHSDTPEIRKYKRFNSEILCAALWGVNLLVGTENGMLLDRS	1074
Qy	1027	GQ	KVYGLIGRRFOQMDVLEGLNLLITISGKNKLRVYILSWLKNKILHNDPEVEKQO	1086
Db	1075	GQ	KYNLINRRRFOQMDVLEGLNVLVTISGKNKLRVYILSWLKNRILHNDPEVEKQO	1134
Qy	1087	WT	VGDMGCGHYRVVYKERIKFLVIALKSSVEVYAWAPKPYHKFMAFKSFADLPHRPLL	1146
Db	1135	WT	VGDLGCGHYRVVYKERIKFLVIALKNAVEIYAWAPKPYHKFMAFKSFADLQHKPLL	1194
Qy	1147	VDL	TVEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPNTDGMEM	1206
Db	1195	VDL	TVEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIVILPKTDGMEM	1254
Qy	1207	LL	CYEDEGVYNTYGRITKDVVLQWGEPTSVAYICSNQIMGWGEKAEIRSVETGHLDG	1266
Db	1255	LVC	YEDEGVYNTYGRITKDVVLQWGEPTSVAYIHSNQIMGWGEKAEIRSVETGHLDG	1314
Qy	1267	VFM	KRAQRLKFLCERNDKVFPASVRSQSSQVYFMTLNRNCIMNW	1312
Db	1315	VFM	KRAQRLKFLCERNDKVFPASVRSQSSQVYFMTLNRNSMMNW	1360

Search completed: August 28, 2004, 00:57:04
Job time : 167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:50:59 ; Search time 157 Seconds
(without alignments)
2636.687 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAPARSLDDILSALRD.....SGGSSQVYFMTLNRNCIMNW 1312

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertibrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6701	96.7	1334	11 Q7TT13	Q7tt13 mus musculus
2	4189.5	60.5	1303	13 Q7SY42	Q7sy42 brachydanio
3	2805.5	40.5	1550	5 Q9W002	Q9w002 drosophila
4	2698	38.9	1082	5 Q8T8M3	Q8t8m3 caenorhabdi
5	2696.5	38.9	1087	5 Q9XYC3	Q9xyc3 caenorhabdi
6	2692	38.9	1096	5 Q95Z17	Q95z17 caenorhabdi
7	2661	38.4	1072	5 Q814B5	Q814b5 caenorhabdi
8	1855	26.8	538	4 Q724L4	Q724l4 homo sapien
9	1826	26.4	1582	4 Q722Y5	Q722y5 homo sapien
10	1702.5	24.6	1455	11 Q9R0G8	Q9r0g8 mus musculus
11	1701.5	24.6	1455	11 Q9R0S4	Q9r0s4 mus musculus
12	1551	22.4	292	4 Q9HBM9	Q9hbm9 homo sapien
13	1458	21.0	792	5 Q9UAN7	Q9uan7 caenorhabdi
14	1377	19.9	728	5 Q9XYC4	Q9xyc4 caenorhabdi
15	993	14.3	334	11 Q9C9S9	Q9c9s9 mus musculus
16	829.5	12.0	169	11 Q61155	Q61155 mus musculus

17	811.5	11.7	1218	5	Q9V8R6	Q9v8r6 drosophila
18	804.5	11.6	1615	4	Q9NV88	Q9nv88 homo sapien
19	804.5	11.6	1616	4	Q8NEV4	Q8nev4 homo sapien
20	800.5	11.6	1613	11	Q8K3H5	Q8k3h5 mus musculus
21	795	11.5	1775	13	Q90XG6	Q90xg6 brachydanio
22	779	11.2	1838	13	Q9DG88	Q9dgg8 morone saxa
23	765	11.0	1310	13	Q800Q6	Q800q6 morone saxa
24	764.5	11.0	1192	4	Q81X64	Q81x64 homo sapien
25	764.5	11.0	1251	4	Q81X68	Q81x68 homo sapien
26	764.5	11.0	1275	4	Q81X65	Q81x65 homo sapien
27	764.5	11.0	1278	4	Q8WXR4	Q8wxr4 homo sapien
28	764.5	11.0	1314	4	Q81X66	Q81x66 homo sapien
29	764.5	11.0	1341	4	Q81X67	Q81x67 homo sapien
30	761.5	11.0	1113	4	Q96N94	Q96n94 homo sapien
31	734	10.6	737	5	Q861X1	Q86ix1 dictyosteli
32	734	10.6	829	3	Q87LH9	Q87lh9 neurospora
33	720	10.4	825	10	Q8SAB1	Q8sae1 triticum mo
34	718	10.4	825	10	Q9ARL7	Q9arl7 hordeum vul
35	713.5	10.3	487	11	Q9J111	Q9ji11 mus musculus
36	712	10.3	950	13	Q9YHC9	Q9yhc9 xenopus lae
37	711.5	10.3	842	10	Q9FNU3	Q9fnu3 oryza sativ
38	710	10.2	836	10	Q24527	Q24527 arabidopsis
39	707.5	10.2	1120	10	Q9LQA1	Q9lqa1 arabidopsis
40	703	10.1	491	13	Q802A6	Q802a6 squalus aca
41	701.5	10.1	497	11	Q9J110	Q9ji10 mus musculus
42	701.5	10.1	539	11	Q80UG4	Q80ug4 mus musculus
43	701	10.1	947	5	Q8SYA1	Q8syal drosophila
44	700	10.1	651	5	Q95ZN6	Q95zn6 caenorhabdi
45	700	10.1	947	5	Q8MLI8	Q8mli8 drosophila

ALIGNMENTS

RESULT 1

Q7TT13 ID Q7TT13 PRELIMINARY; PRT; 1334 AA.

AC Q7TT13; 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Richards S., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052474; AAH52474.1; -
KW Hypothetical protein.
SQ SEQUENCE 1334 AA; 150166 MW; CB462D07CFBDE8F5 CRC64;

Query Match 96.7%; Score 6701; DB 11; Length 1334;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 1277; Conserved 18; Mismatches 14; Indels 28; Gaps 5;

QY 1 MGDPAARSLLDIDLSALRDPAGIFELVEVNGTGYGVYKGRHVKTGLAIAKMDVTE 60
DB 1 MGDPAARSLLDIDLSALRDPAGIFELVEVNGTGYGVYKGRHVKTGLAIAKMDVTE 60

QY 61 DEEEIKQBINMLKYSHERNTATYGAFFIKSPGNDQDLWMEFCGAGSVTLVKNT 120
DB 61 DEEEIKQBINMLKYSHERNTATYGAFFIKSPGNDQDLWMEFCGAGSVTLVKNT 120

QY 121 KGNALKEDCIAYICREILAGLAHLAHKVIHRDIKQNVLLTENAELVDFGSAQLDR 180
DB 121 KGNALKEDCIAYICREILAGLAHLAHKVIHRDIKQNVLLTENAELVDFGSAQLDR 180

QY 181 TVGRNTFTGTYPWAPEVIACDENPDATYDYSOIWSLIGITAIEMAGAPPLCDMPWR 240
DB 181 TVGRNTFTGTYPWAPEVIACDENPDATYDYSOIWSLIGITAIEMAGAPPLCDMPWR 240

QY 241 ALFLIPRNPPRLKSKWKKFDIDTCLIKTYLSRPTEQLAKFPPIRDOPTQVR 300
DB 241 ALFLIPRNPPRLKSKWKKFDIDTCLIKTYLSRPTEQLAKFPPIRDOPTQVR 300

QY 301 QLKDHIDRSRKRGEKEETEYYSYSGSEEDDGHGEGEPSSIMNVPGESTLRREFLR 360
DB 301 QLKDHIDRSRKRGEKEETEYYSYSGSEEDDGHGEGEPSSIMNVPGESTLRREFLR 360

QY 361 ENKSNSEALKQO 420
DB 361 ENKSNSEALKQO 420

QY 421 LQEKQORLEDMOALRREERERQAEQYKQKLEQORQSERLQOQOQOQOQOQOQO 480
DB 421 LQEKQORLEDMOALRREERERQAEQYKQKLEQORQSERLQOQOQOQOQOQO 480

QY 481 QO 538
DB 481 QO 538

QY 539 KSKPGSTGPEPPIPOASPGPGPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLP 597
DB 539 KSKPGSTGPEPPIPOASPGPGPGLPGLPGLPGLPGLPGLPGLPGLPGLPGLP 597

QY 581 --SLQDQPTRNLAAPFASHPDPP-ALPAPTATPSARGAVIRONSPTSEGPSPNPPAW 637
DB 581 --SLQDQPTRNLAAPFASHPDPP-ALPAPTATPSARGAVIRONSPTSEGPSPNPPAW 637

QY 598 SOSLQDQPTRNLAAPFASHPDPPAAVPTTATPSARGAVIRONSPTSEGPSPNPPSW 657
DB 598 SOSLQDQPTRNLAAPFASHPDPPAAVPTTATPSARGAVIRONSPTSEGPSPNPPSW 657

QY 638 VRPDNEAPKPVORTSSATLNTSGAGSRPAQAVRAPRSNSAWQYLQRAERGT 697
DB 638 VRPDNEAPKPVORTSSATLNTSGAGSRPAQAVRAPRSNSAWQYLQRAERGT 697

QY 658 VRPDNEAPKPVORTSSATLNTSGAGSRPAQAVRAPRSNSAWQYLQRAERGT 717
DB 658 VRPDNEAPKPVORTSSATLNTSGAGSRPAQAVRAPRSNSAWQYLQRAERGT 717

QY 698 PGPPPAQPPGNASNDPLRSDPGWERSDVLPAHIGHLPQAGSLR--RNRVQASKL 755
DB 698 PGPPPAQPPGNASNDPLRSDPGWERSDVLPAHIGHLPQAGSLR--RNRVQASKL 755

QY 718 PGPPPAQPPGNASNDPLRSDPGWERSDVLPAHIGHLPQAGSLRNRNRVQASTKL 777
DB 718 PGPPPAQPPGNASNDPLRSDPGWERSDVLPAHIGHLPQAGSLRNRNRVQASTKL 777

QY 756 DSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSEEEVESSED 815
DB 756 DSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSEEEVESSED 815

QY 778 DSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSEEEVESSEE 837
DB 778 DSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSEEEVESSEE 837

QY 816 DEEEGEGPAGSRDTPGSRSDGTDTSVSTWVHDVBEITGTQPPYGGGTWVQRTPEE 875
DB 816 DEEEGEGPAGSRDTPGSRSDGTDTSVSTWVHDVBEITGTQPPYGGGTWVQRTPEE 875

QY 838 BEEEGDGEPSGRSDTPGSRSDGTDTSVSTWVHDVBEISGTQPSYGGGTWVQRTPEE 897
DB 838 BEEEGDGEPSGRSDTPGSRSDGTDTSVSTWVHDVBEISGTQPSYGGGTWVQRTPEE 897

QY 876 RNLHADSNYTNLDDVQPSHSPTENSKGQSPGSGDGYQSRGLVAKPKSGSFTWVF 935
DB 876 RNLHADSNYTNLDDVQPSHSPTENSKGQSPGSGDGYQSRGLVAKPKSGSFTWVF 935

QY 898 KSLLLADSNYTNLDDVQPSHSPTENSKGQSPGSGDGYQSRGLVAKPKSGSFTWVF 957
DB 898 KSLLLADSNYTNLDDVQPSHSPTENSKGQSPGSGDGYQSRGLVAKPKSGSFTWVF 957

QY 936 DLGIYQPGSGDSIPITALVGGEGTRLDQLQYDVRKGSVNVNPTNTRAHSETPEIRKYK 995
DB 936 DLGIYQPGSGDSIPITALVGGEGTRLDQLQYDVRKGSVNVNPTNTRAHSETPEIRKYK 995

QY 958 DLGIYQPGSGDSIPITALVGGEGTRLDQLQYDVRKGSVNVNPTNTRAHSETPEIRKYK 1017
DB 958 DLGIYQPGSGDSIPITALVGGEGTRLDQLQYDVRKGSVNVNPTNTRAHSETPEIRKYK 1017

QY 996 KRFNSEILCAALWGVNLLVGTENGMLLDRSGQGGKQYVGLIGRRRFPQOQMDVLEGLNLLITI 1055
DB 996 KRFNSEILCAALWGVNLLVGTENGMLLDRSGQGGKQYVGLIGRRRFPQOQMDVLEGLNLLITI 1055

QY 1018 KRFNSEILCAALWGVNLLVGTENGMLLDRSGQGGKQYVGLIGRRRFPQOQMDVLEGLNLLITI 1077
DB 1018 KRFNSEILCAALWGVNLLVGTENGMLLDRSGQGGKQYVGLIGRRRFPQOQMDVLEGLNLLITI 1077

QY 1056 SGKRNKLVYVLSWLNKILHNDPEVEKKQGWTTVDMEGCGHRYRVYKRIKFLVIALK 1115
DB 1056 SGKRNKLVYVLSWLNKILHNDPEVEKKQGWTTVDMEGCGHRYRVYKRIKFLVIALK 1115

QY 1078 SGKRNKLVYVLSWLNKILHNDPEVEKKQGWTTVDMEGCGHRYRVYKRIKFLVIALK 1137
DB 1078 SGKRNKLVYVLSWLNKILHNDPEVEKKQGWTTVDMEGCGHRYRVYKRIKFLVIALK 1137

QY 1116 SSVEVYAWAPKPYHKWAFKSFADLPHRPLVLDLTVBEGQRLKVTYGSAGFHAVDVDSG 1175
DB 1116 SSVEVYAWAPKPYHKWAFKSFADLPHRPLVLDLTVBEGQRLKVTYGSAGFHAVDVDSG 1175

QY 1138 NYVEVYAWAPKPYHKWAFKSFADLPHRPLVLDLTVBEGQRLKVTYGSAGFHAVDVDSG 1197
DB 1138 NYVEVYAWAPKPYHKWAFKSFADLPHRPLVLDLTVBEGQRLKVTYGSAGFHAVDVDSG 1197

QY 1176 NSYDIYIPVHIQSQITPHAIIFLENTDGMELLCYEDGEGVYVNTYGRILKDVVLQWGE 1235
DB 1176 NSYDIYIPVHIQSQITPHAIIFLENTDGMELLCYEDGEGVYVNTYGRILKDVVLQWGE 1235

QY 1198 NSYDIYIPVHIQSQITPHAIIFLENTDGMELLCYEDGEGVYVNTYGRILKDVVLQWGE 1257
DB 1198 NSYDIYIPVHIQSQITPHAIIFLENTDGMELLCYEDGEGVYVNTYGRILKDVVLQWGE 1257

QY 1236 TSVAYICSNQIMGEBKAIBRSVETGHLGDFVHHKRAQLKFLCERNDKVFFASVRSGG 1295
DB 1236 TSVAYICSNQIMGEBKAIBRSVETGHLGDFVHHKRAQLKFLCERNDKVFFASVRSGG 1295

QY 1258 TSVAYICSNQIMGEBKAIBRSVETGHLGDFVHHKRAQLKFLCERNDKVFFASVRSGG 1317
DB 1258 TSVAYICSNQIMGEBKAIBRSVETGHLGDFVHHKRAQLKFLCERNDKVFFASVRSGG 1317

QY 1296 SSQVYFMTLNRNCIMNW 1312
DB 1318 SSQVYFMTLNRNCIMNW 1334

RESULT 2
Q7SY42
ID Q7SY42 PRELIMINARY; PRT; 1303 AA.
AC Q7SY42;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
EMBL; BC055134; AAH55134.1; -.

[illegible]


```
QY 819 EGGGPAEGSRDTPGSGDGTSDSVTWVHVDVEITGTPPYGGTVMVQRTPEERNL 878
Db 635 EG-----NEPLMFKEI 645
QY 879 LHADNGYTNLPDVVOPSHSPENSKGQPPSKDGGDYQSRGLVKAPEKGSF----- 931
Db 646 NASSSRG--ALPDLAPKSPQLRRQINDQTRQMSDDRADEQPNFGQNSDSRSIQHSFSNR 703
QY 932 ---TWFDVLGIYQPGSGSDSIPITALVGEGT-----RLDLOLYDVRKGSVVNVNPT 980
Db 704 DREKSFVG---YFGGAG-----AGGTVNRPGRPDQINQVQ-----VNVTPN 743
QY 981 N--TRAHSTPIRKYKRFNFSEILCAALGWVNLVGTENGMLLDRSGQGVYGLIGRR 1038
Db 744 SNGTPAENDAPRIKYKRFNFSEILCAALGWVNLVGTENGMLLDRSGQGVYGLISRR 803
QY 1039 RFQQMDVLEGLNLLTISGRNKLRYVYLSWLNKILHND-----PEVEKKQGTWTVGDM 1093
Db 804 REDQMTVLEGGQNILATISGRKRIIRVYVLSWLNKILRTEGAGSANTTEKRWGNVNVGDL 863
QY 1094 ECGHYRVVYKRIKFLVIALKSSVEVVAWAPKPYHKEMAFKSFADLRPLLDVLTVEE 1153
Db 864 QGAIHFIVRYEKIFLVGLLESSIETVAMAPKPYHKFMFSFGSLSHVPLDIVLTVED 923
QY 1154 GQRLKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSIQITPHAIIFLENTDGMELLCYDE 1213
Db 924 NARLKVLYSGTGFAHDLSDAAVYDIYTPAGSGOTTTPHCIVVLPNSNGMQLLCYDNE 983
QY 1214 GYVNTYGRILIKDVVLQWGMETPSVAYICSNQIMGKEKAIIRSVETGHLDGVPMHKRA 1273
Db 984 GYVNTYGRMTNVVLQWGMETPSVAYISTGQIMGNKAIIRSVDTGHLDGVPMHKKA 1043
QY 1274 QRLKFLCERNKDVFFASVRSVGSSGVYFMTLNRNCIMNW 1312
Db 1044 QKLKFLCERNKDVFFSSAGGSGCQIYFMTLNRKGLTNW 1082

RESULT 5
QYXVC3
ID Q9XYC3 PRELIMINARY; PRT; 1087 AA.
AC Q9XYC3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase MIG-15.
GN ZC504.4 OR MIG-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Zhu X., Acharya P., Hedgecock E.;
RT "MIG-15, a NIK ortholog of the STR20 family of serine/threonine
RT protein kinases, is involved in cell migration and signal transduction
RT in C. elegans.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF087131; AAD14593.1; -.
DR EMBL; Z50029; CAB63416.1; -.
DR WormPep; ZC504.4a; CE25672.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
```

```
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01071; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00106; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1087 AA; 121543 MW; 49FC2C39893F95CD CRC64;

Query Match 38.9%; Score 2696.5; DB 5; Length 1087;
Best Local Similarity 43.7%; Pred. No. Se=147;
Matches 596; Conservative 157; Mismatches 268; Indels 343; Gaps 34;

QY 10 LDDIDLALRDPAGIFELVEVVGNGTYGVYKGRHVKTGLAAIKVMDVTDEDEEEIKQE 69
Db 6 LDEIDLNSLRDPAGIFELIEVVGNGTYGVYKGRHVKTGLAAIKIMNINEDEDEEIKLE 65
QY 70 INMLKKYSHRNINATYYGAFIKKSP--PGNDDQLWLMVEFCGAGSVTDLVKNYTKNALKE 127
Db 66 INMLKSHRNINATYYGAFIKKLPSTGTGHDQLWLMVEFCGSGITDLVKNYTKGSLKE 125
QY 128 DCIAYICREILRGLAHHLHAHKVIRHDIKQGNVLLTENAELVDFGVSAQLDRTVGRNT 187
Db 126 EWIAIYICREILRGLYHLHQSIVIRHDIKQGNVLLTDSAEVLDVDFGVSAQLDRTVGRNT 185
QY 188 FIGPYWMAPEVIACDENPDATYDYSIDWSLGTATJEMAEAGAPLCDMHPMALFLIPR 247
Db 186 FIGPYWMAPEVIACDESPATYDSRDLWSLGTALJEMAEAGHPPLCDMHPMALFLIPR 245
QY 248 NPPRLK-SKWSKKFIDFTCLIKTYLSRPPTEQLLKPPPIRDQTEROVRQLKQHI 306
Db 246 NPPPKLRNKKWKTKFTFETVLVDYHORPYTGALLRHPFIKEQHEQTHIRSIKEHI 305
QY 307 DRSRKKEKEEYEYSGSEEDDSHGEGEPSSIMN-----VPGESTLRREFLRILQ 359
Db 306 DNRN--RVKDDADYEYSGSEDEDEPSNNRGSPNGIRDSSESSMIAMDNTLRGFKQLQ 363
QY 360 QENKSNSEALKQ--QQLQQQQQRRDPEAHIKHLHQRRRIEEQKE-----ERRRVE 409
Db 364 ESSRGFAEPGAQQLRRLPQPAPAP-----FOYQSRVYEPRESSEVKLRVSSRGAA 417
QY 410 EQQREREQRKLQEKQORLED-----MQALRREEREERQAEQEYKKEQLEEQRS 462
Db 418 DGPRHSPASRPVSVHHQSPQOSHAPAPHLADLANYEKRRRSRE----- 463
QY 463 ERLQRLQQLQEHAYLKSLLQQQQQQQLLQKQQQLLPQGRKPLVHYGRMNPADKPWARE 522
Db 464 ERRERERQAHAM----- 476
QY 523 VEERTRNVKQNSPLAKSPGSGTGPBPPIQASPGPGPLSQTPPMQRP-----VEQEG 577
Db 477 -----PIARVSASVAP--QQSRKMSPELLITHVKPED- 507
QY 578 PHKSLQDQPTNLAAFPASHDDPDAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPPAW 637
Db 508 -----LDVLASELSKMGHH-----NGRSRE-----SMSPP-- 534
QY 638 VRPDNEAPPKVPQRTSSITATALTNSGAGSRPAQAVRARPRNSAWQIY--LQRRABRG 695
Db 535 -----PPAPPPEASISITDIDVG-----ELDNGADAEDDLKIMNGE-GT 578
QY 696 KPDPGPPAQPDPGNASSN--PDLRRSDPGWERSDVLDPASHGHLPAQGLSERNVRVASS 753
Db 579 LRGNKPL-PPTPDGENTLVSDVRNGNG-----NSGH-----GAYKKKI----- 619
QY 754 KLDSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSEEVSS 813
Db 620 -----PEIRPGITISLDD----- 634
```


Db 621 SNEEG-----NEPLM 631
Qy 875 ERLNHLADSNYTNLPDVVQPSHSPTEKSGSPKSGDGYQSRGLVKAPKGSF--- 931
Db 632 PKPINASSRG--ALPDLPLKSPQLRRQINDQTRQMSDDRADEQPNQFQNSDRSSIQHS 689
Qy 932 -----TMFVDLGIYQPGSGDSIPITALVGEET-----RLDQLOYDVKSGSVN 976
Db 690 FSNRDREKSFVG---YFGGAG-----AGGGTVNRPGRPDINQVQ-----VN 729
Qy 977 VNPNTN--TRAHSETPEIRYKRFNSEILCAALWGNLLVGTENGLMLLDRSGQGGKQVGL 1034
Db 730 VTPNSNGTPAENDAEIRYKFKFGEILCAALWGNLLIGTDSGLMLLDRSGQGGKQVPL 789
Qy 1035 IGRBPQMDVLEGNLLTISGRKNKLRVYVLSLNRKILND-----PEVEKKGQWTT 1089
Db 790 ISRRFPQMDVLEGNLLTISGRKNKLRVYVLSLNRKILND-----PEVEKKGQWTT 1089
Qy 1090 VCDMEGCGHYRVVYKRIKFLVIALKSSVEVYAWAPKPKHMKAPKSPADLPRLVLDL 1149
Db 850 VGLQGAHFKIVRYERIKFLVGLGLESSIEIYAWAPKPKHMKAPKSPADLPRLVLDL 909
Qy 1150 TVEGRQRLKVIYSSAGFHAVDVDSGNSYDIYIPVHIQSIPTPHAIIFLPNTDGMELLC 1209
Db 910 TVEDNARLKVLGSGTGGFHAIDLSAAVDIYTPAQSGQTTPHICIVLVPNSGQMLLC 969
Qy 1210 YEDEGVYNTYGRRIIKDVVLWNGEMPTSVAYICSNQIMGWGEKAIEIRSVETGHLDGVM 1269
Db 970 YDNEGVYNTYGRMTKNNVVLWNGEMPTSVAYICSNQIMGWGEKAIEIRSVETGHLDGVM 1029
Qy 1270 HKRAQLKFLCBRNKDKVPFASVRSGSSQVYFMTLNRNCIMW 1312
Db 1030 HKKAQKLKFLCBRNKDKVPFASVRSGSSQVYFMTLNRNCIMW 1072
RESULT 8
Q724L4
ID Q724L4 PRELIMINARY; PRT; 538 AA.
AC Q724L4
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh P.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schumacher J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055427; AAH55427.1; -.
KW Hypothetical protein.
SQ SEQUENCE 538 AA; 60315 MW; 598FBCA4434C5F70 CRC64;
Query Match 26.8%; Score 1855; DB 4; Length 538;
Best Local Similarity 69.2%; Pred. No. 6.7e-99;
Matches 375; Conservative 46; Mismatches 89; Indels 32; Gaps 10;
Qy 798 KKANDYSSSSSEVSESSDEDEEGGPAEGS---RDTF-----GGRSDGDTDSVSTMVVH 849
Db 2 KKVDYSSSSSESSSESESESESETHDGTAVASDIPRLIPTGAPGNEQVNVGVGTH 61
Qy 850 DVEITGTQPPYGG-----GTVMVQRTPEERNLHADSNGYT---NLPDVVQPSHS--- 898
Db 62 GLE--TSHADSPSGSISREGTLMIRETSGEKKRSHSDSNGFAGHINLPDLVQSSHSPAG 119
Qy 899 -PTNSKQSQSPSKD--GSGDYQSRGLVKAPOKSFMTFVDLGIYQPGGS-----GDSIPT 951
Db 120 TPTTEGLGRVSTHSDQEMDSGTGYGMG---SSTKASFTPFVDPRVYQTSPTDEDEDEESSA 176
Qy 952 TALVGGECTRLDQY--DVRKGSVNVNPTNTRAHSETPEIRYKRFNSEILCAALWGV 1010
Db 177 TALFTSELLRQEQAKNEARKISVVNVNPTNIRPHSDTPEIRYKRFNSEILCAALWGV 236
Qy 1011 NLLVGTENGLMLLDRSGQGGKQVYGLIGRRRFOQMDVLEGNLLTISGRKNKLRVYVLSWL 1070
Db 237 NLLVGTENGLMLLDRSGQGGKQVYGLIGRRRFOQMDVLEGNLLTISGRKNKLRVYVLSWL 296
Qy 1071 RNKILNDPEVEKKGQWTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEVYAWAPKPHK 1130
Db 297 RNRILNDPEVEKKGQWTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEVYAWAPKPHK 356
Qy 1131 FMAKSFADLPRLVLDLTVEEGRQKVIYSSAGFHAVDVDSGNSYDIYIPVHIQSOI 1190
Db 357 FMAKSFADLPRLVLDLTVEEGRQKVIYSSAGFHAVDVDSGNSYDIYIPVHIQSOI 416
Qy 1191 TPHAIIIFLPNTDGMELLCYEDEGVYNTYGRRIIKDVVLWNGEMPTSVAYICSNQIMGW 1250
Db 417 TPHAIIIFLPNTDGMELLCYEDEGVYNTYGRRIIKDVVLWNGEMPTSVAYICSNQIMGW 476
Qy 1251 EKAIEIRSVETGHLDGVMFHMRAQLKFLCBRNKDKVPFASVRSGSSQVYFMTLNRNCIM 1310
Db 477 EKAIEIRSVETGHLDGVMFHMRAQLKFLCBRNKDKVPFASVRSGSSQVYFMTLNRNCIM 536
Qy 1311 NW 1312
Db 537 NW 538
RESULT 9
Q722Y5
ID Q722Y5 PRELIMINARY; PRT; 1582 AA.
AC Q722Y5
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein DXFP686A17109.
GN DXFP686A17109.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human fetal kidney;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538345; CAD98108.1; -.
KW Hypothetical protein.

SQ	SEQUENCE	1582 AA; 178289 MW; B1002851574E8079 CRC64;
	Query Match	26.4%; Score 1826; DB 4; Length 1582;
	Best Local Similarity	31.2%; Pred. No. 1.3e-96;
	Matches	503; Conservative 232; Mismatches 529; Indels 350; Gaps 43
Qy	1	MGDPAPARSLDDIDLALSRLDPAGIFELVEVGNNGTYGVYKGRHVKTGQLAAIKWMDYTE 60
Dd	1	MAGPGWRDREVTDLGLHLPDTGISLTKTIGLTGYRIYLGHEKTAFAVKVMNARK 60
Qy	61	-----DEBEIKOEINMLKKYSHRNIATYGAFIKKSPGG 96
Dd	61	TPLPEIGRRVRWNKYQKSVGWRYSDDEEDLRTELNLRLKSYSPHKNVISVGPAFFKLISPPG 120
Qy	97	NDDQLWLMEFCGAGSVDTLVKNTKGNALEDICAIYICREILRGLAHLHAHVIRHDIKG 156
Dd	121	QRHQLMWMELCAAGSVTDVVMTSNQSILEGWIAIYICREILQGLAHLHAHVIRHDIKG 180
Qy	157	QNVLITENAEVKLVDFGVSQAOLDRTVGRNRFTFIGTPYMWAVEVIACDENPDATYDYSDI 216
Dd	181	QNVLITHNAEVKLVDFGVSQAQVSRTRNGRNRSFIGTPYMWAVEVIDCEDPRRSYDRSDV 240
Qy	217	WSLGITALEMAEGAPPLICMHPMRALFLIPNPPLRLSKWKSKKFIDFIDCLLIKTVLS 276
Dd	241	MSVGITALEMAEGAPPLCNLCQLEAFILFRESAPTIVKSGWSRKPFHFMEKTIKNFLF 300
Qy	277	RPPTQELLKFPFIKDQPTERVRIQLKDH----IDRSRKRG-----EKET-EYEYSGSE 327
Dd	301	RPTSANMLQHPFVRDIKNERHVESLTHLGTGIKKROKKGIPLFIREBAIKEQYTVRR 360
Qy	328	EEDDS-----HGEBGP-----SSIMVPGESTLRRBFL 356
Dd	361	FRGPSCTHELLRLPTSSRCRPLRVLHGSPSPRWLPDREEPQVALQLOGAA---RVFM 417
Qy	357	RLQGENKSNEALKQQOOLQOOQBDPAHIKHLLHORQRIEEQKEERRRVEEQORER 416
Dd	418	PLOALD-SAPKPLQGQAQAPORLQGAARVF--LQAVQAKASKPLQMOTKAPPRLR 473
Qy	417	EQRKLQEKBQORRLDMQALRREREORAEQERYKRKOLEEORSERLQRLQLOQEHAYL 476
Dd	474	AARVLMPLQAVRAPRLLQVOVSVKQAQTQTSFPDLOQVPFEFSQDVPEQQRQG 533
Qy	477	KSLQOQQOQQOQLQKQQOQLLPGDRKPLHYHGRGMNPADKPAWA---REVEERTMNKQQ 533
Dd	534	QAPEQQRHNQVPEQELEQNAPEQEV-----QEQAAPAQAGTEABEPESLRVNVAQV 587
Qy	534	NSPLAKSK-----PGSTGPBPPIPOASPGCPGLSOTPPMQRPVPEQGPBKSLQDQPTR 588
Dd	588	FLPLLSSQDHHVLLPHLHDTQVLIIPVEGTEGSPAQAWTLEPPP--QAIGSVQALIEGLSR 645
Qy	589	NLAAFPASHDPPDAIPAPTATPSARGAVIRONSOPTSSEGPCSPNPMPVWPDEAPPKV 648
Dd	646	DLLRAPNNNKPLGPLOTLMENLSSN--RYSQPEQA-----REKSKV 688
Qy	649	PORTSSIALTNLS--GAGGS-RP-----AOAVRARPRSNAWO-IYLQRAE-RGTPKP 698
Dd	689	STLQALAKRLSPKRFGAKSWREPKELSLDLEARRQRRQREDIFNQHEELRQVDKD 748
Qy	699	PGPPA-----QPFGPNASNPDLRSDPGWERSDVLPSHGH-----737
Dd	749	KEDESSONDEVFHSIQAEVQIEPLK--YISNP---KKIEVQERSPSV-PNNQDHAHVK 802
Qy	738	-----LPQAGSLF-----RNR-----VCASSKLDSSPVLSGNKAKP-----769
Dd	803	FSSSVPORSLLEQAQKPIDIRQSSQNRQNWLAASESSSESPTVGRSSOSPYPYSTID 862
Qy	770	-----DDH-----RSSRP-----GRPADFLLKERTLDEAP--RPPKKAMDY--803
Dd	863	QKLLVDLHVPDGFKVGKISPPVILTNEWGVNALUSEIFRDNLIPAPVIPPEEDGDYVE 922
Qy	804	-SSSESEVESSEDDE-----818
Dd	923	LYDASADTGDDGDDDESNDTFDYDHANGDDLNOVDANDVCKDDHDDNNKFNVDVNN 982

Qy	819	-----EEGGPAESR-----DTPGSRDGTDSVSTVWVHD	850
Db	983	NYEAPSCPRASYCRDGSCKDQGDYDSRGKEEAYRGYGHSTANRSHGSSAASDAAIGD	1042
Qy	851	VEEITGT-----OPPYGGGTVMVQRTPEBERNL-LHADSN-GYTNLPDVVQPSHSPTEN-	902
Db	1043	QEEHANIGSERGSEGDGKGVVFTSESGALNGEENCSETDQGLKRPASQDFEYL	1102
Qy	903	-----SKGQSPPSKXGSDGYOSRGLVKAPGKSSFTFMFVDLIGIYQPGSGSDS	948
Db	1103	QEEPPGGNEASNAIDSGAAPSAFDHESD--NKDISESPQTQDSFSAHSSPSKSGKMSADA	1160
Qy	949	IPITALVGGEGTLDQLQYDVRKGSVNVNPTTRAHSETPETRKVKKPFNSFILLCAALW	1008
Db	1161	NFASAILYAGFVEVP--EESPQKQPSVNVNPLVSPACKPLIHMYEKEFTSEICGSLW	1218
Qy	1009	GVNLLVGTENGLMLDRSGGKGYGLIGRRRFQOMDVLEGINLLITISGRKNKLRVVYLS	1068
Db	1219	GVNLLIGTRSNLYIMDRSGKADTKLIRRPFFQIQVLEPFLNLLITISGHQWRLRVYHLT	1278
Qy	1069	WLNRKILHNDPEYKKQ-----GWTTVGDMEGCGHYRVVKYRIKFLVIAIALKSSVEV	1120
Db	1279	WLNRKILNNDPESKRQEEMLKTEEACKAIDKLTCGEHFSVLQHEETTVIAIALKSSIHL	1338
Qy	1121	YAWAPK-----PYHKF-----M	1132
Db	1339	YAWAPKSFDEBTAIKVICDSADSEGYMSQYAIRILAKIQAADVNNRFPKRPDELLHL	1398
Qy	1133	AFKSFADLPHRPLLVLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDIVPVHIQSQITP	1192
Db	1399	KLKVFPTLDHKPTVDLATGSEKRLKIFFSSADGYHLIDAESEVMSDVLTKNPLEIIP	1458
Qy	1193	HAIFIPTNDQEMBLICYEDEGVVNTYGRIIKDVVLQWCEMPTSVAYICSNQIMGWGEK	1252
Db	1459	QNIITILPDCILGIMLTFNAEALSVEANEQLFKILEMWDIPSSIAFECTQRTTGWGQK	1518
Qy	1253	AIERSVETCHLDGVFMFKRAQRLKELCERNDKVFASVRSGGSSOVYFWTLNR	1306
Db	1519	AIERSVLSQSVLESELKRRSIKKLRLCTRGDKLFFFTLNRNHSRVYFWTLGK	1572
RESULT 10			
Q9R0G8		PRELIMINARY;	PRT; 1455 AA.
ID	Q9R0G8		
AC	Q9R0G8;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Nck-interacting kinase-like embryo specific kinase.		
GN	NRK OR NESK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20347139; PubMed=10801798;		
RA	Nakano K., Yamauchi J., Nakagawa K., Itoh H., Kitamura N.;		
RT	"NESK, a member of the germinal center kinase family that activates		
RT	the c-Jun N-terminal kinase pathway and is expressed during the late		
RT	stages of embryogenesis";		
RL	J. Biol. Chem. 275:20533-20539(2000).		
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; AB035267; BAA87066.1; -.		
DR	MGD; MGI:1351326; Nrk.		
DR	GO; GO:0007256; P:activation of JUNKK; IDA.		
DR	InterPro; IPR001180; Citron.		
DR	InterPro; IPR00719; Prot_kinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR008271; Ser_thr_pkin_AS.		
DR	Pfam; PF00780; CNH; 1.		
DR	Pfam; PF00069; pkinase; 1.		

```
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1455 AA; 163675 MW; C9B24D1F207775AD CRC64;

Query Match      24.6%; Score 1702.5; DB 11; Length 1455;
Best Local Similarity 30.6%; Pred. No. 1.6e-89;
Matches 474; Conservative 231; Mismatches 497; Indels 347; Gaps 46;

QY 1 MGDPAASLDDIDLSALRDPAGIFELVEVNGNGYGVQYKGRVKTGQLAAIKMDYTE 60
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1 MAGPGSRDKVETDLGQLPDPGTGIFSLDKAIGLGTGRIFLGIHEKTSGLVAVKMSARK 60
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 -----DEEEIKQEIIMLKYSHRNIATYVYAFIKKSPG 96
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 TPLPEIGRRVRVNYKQSVGRWYSDEEDRLTELNLRLKYSFHKVIVTFYGAFFKLNPPG 120
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 97 NDDQLWLMVFCAGSVTDLVNTKGNALKEPCIAIYICREILRGSLAHKVIHRDIKG 156
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 HQHQLWVWMLCAAGSVTDVVRMTNRQSLKEDWIAIYICREILQGLAHKVIHRDIKG 180
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 157 QNVLLTENAELVKLDFGSAQLDRVTGRNTFICTPYWMAPEVIACDENPATYDYSDI 216
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 QNVLLTHDAEKVIVDFGSAQVSRVTNGRNSPTGTPYMAPEVIHCDPRCSYDYSRV 240
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 217 WSLGATATAMAGAPLCPDMHPRALFLIPRPPRLSKKWKPFIDFIDTCLIKTVLS 276
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 WSVGITATAMAGAPLCKQPLEALCVILREAPKVASGWSRKFQNFMCMLKNFLF 300
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 277 RPPTQLLKFPIRDPQTOVRQIQLKDHIDRSRKRGKEKETETEYSGSEEDDSH--- 333
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 RPTSGNMLLPVPHDIKNRRVVESLTKHLTGIIQKR-EKKGI PVAPEGEAAKEQYITR 359
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 334 ---GEGEPSSIMNVPGSSTLRREFLRLOENKNSKALKQOQLOQO---QORDPEAHK 388
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 RFRGSPCTP-ELLRVFTSSRCRP--LRVLHGEPQPRWLPDQEDPDQELQQLAKAGVF 416
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 389 HLLHQKQ-----RRIEQQEER-----RVEEQRREREQKLEKEQORLE 431
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 417 MPLHSQDNTSKLFPKQVAVPYLRGAQVMPVLQVEAPPQVSAQAQMLSLPTQ----- 472
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 432 DMQALRBEERQAREEQ-EYKRKQLEBQROSERLQRLQ-QEAYLKSLOOQOQOQOQL 489
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 473 DNKATSPVQAPVAGSQQAQHEALETEQKOLDQVPEEFQGDRA-----PEQPRQOQ 525
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 KQOQOQLLPGRKPLHYHGRGNPKADKPAWAREVEERTMKNQNSPLAKSPGSGTGP 549
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 AAEQOQI---HNVP-----PRQPEEDREPEQAEVQBEAVEFPQAEIE---DKEP 569
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 550 PIPQASPPGP-LSQTPPMQRPV-----EPQEGPHKSLQD-QPTNLAAFP 595
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 570 EVVQVHAQVLLPLSQNRHLLPLHLDRQLLIPVGEQNEEVPRAQWLEASRAGVQA 629
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 -----SHPDPAIPAPTA---TPSARGAVIRONSPTSEGPGSPNPAPWVRPDNEAPPK 647
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 630 LIEGLSRD---LLRAPNAVTKPLGLQIFLENL---STDGFYTEPEP----- 671
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 648 VPQRTSSTATLNTSGGSRP-----AQAV-----RAPRNSAWQ-IYLO-- 688
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 672 -TQKKKSVASLKAIAKRLAPKPRAKALWLEDFFESDVETSRRRRRHRWEDIFNQHE 730
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 699 ---RA-----ERKTPKP----- 699
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 731 EQLREVRNDRDSSDNDEVFHSIOAEVQIIEPHAAANPAGNEVHESAPMPCNRRNTRVK 790
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 700 -GPPAQPPGPNASNPDLRSDPCWERSDVLPSHGHLPOAGSLERNVGCASKLDS 758
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 731 FSPSVGEEPSLEEAQPOQOQOQOQPNVIRPNCLNPNQFQAQSDSSSEEDSPVTRKSS 850
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

759 P-----VLSP-----GNKAKPD---DHRSPGRPADPVL 784
851 PPYSTIOKLLIDIHVPDGFVKVIGIPVYLTNFWGVNALSSEIFWDDWIWPTPA----- 906
785 LKERTLDEAPRPKKAMDYSSSESESESEDEEGEGGPAEGSRDTPGGSDGDTDSVS 844
907 -----RPPEDGDY-VELYDADANANGDEEVANGAYEDPRDGGANGHDD----- 948
845 TMVVDHVEEITGTO-----PPYGGGTWVQRTPEERNLLHADSNGYTNLPDVVQPSHSPT 900
949 --MNNQLDQANGYEGHGAAGYNGD-----VGGNHGAAPNGPRANYPRAGIL 993
901 ENSKGQSPPSKDSG---GDVQSRGLVAKPGKSFTMFVDLGIYQPGSGG----- 946
994 KNGHNDGRLNRRGAFGVFGDAAARAFHGAAGEAG-AAFGNHGANGRNGRNRNREANGRN 1052
947 -----DSIPITALVGGEGTRLDQLQDYVRKGSVNVV- 977
1053 EENGAFGRDQHVPEPEHEESDRCTETSDSIAL-----EITSFDGEQNSGRPVSSTTMG 1106
978 -----NPTNTRAUSE-----TPEIRKYKGRFNSILCAALWGVNLLVGTENGLMLDRSG 1027
1107 FPIGRSPRGSDFGSDISYNSPILHVYEKDFSSSEVYCGSLWGVNLLLTQSHLYLMDRSG 1166
1028 QGKYVGLIGRRRFOQMDVLEGLNLLITTSKRNKLRVYVLSWLRNKLHNDPEVEKKO-- 1085
1167 KAEIVKLIKRPFQIQVQLNLLITTSKRNKLRVYVLSWLRNKLHNDPEVEKKO 1226
1086 -----GWTVDGMEGCHYRVVYKIFKFLVIAIKSSVEVYAWAPKPYHKHFAFKSPAD 1139
1227 MRKEEACKAIDKLIGCEHFSVLQHEETTYIAVAKSSIHLFAWAPKSFDENTAIKVPPT 1286
1140 LPHRPLVLTVEGQRLKVYIGSSAGFAHVDVDSGNSYDIYIPVHIOSQIT--PHALIF 1197
1287 RDLKPLTVDLAVGSEKTKLIPFSSANGYHIIAESA-----EVMSEVTLPNNNVVI 1336
1198 LPNTDGMELLCYEDEGVYVNTYGRITIKDVVLQNGEMPTSVAYICSNQICMGKAIIR 1257
1337 LPDCLGLGVMLSLNAEASAEANQLKILDVWKDIPSSVAFECTKRITGWDQKAIER 1396
1258 SVETGHLDGVMFKRAORLKFCLERNDKVFPFASVRSGSSQVYFMTLNR 1306
1397 SLQSTILENELKRSIKKRLFLCARGDMFFASTLSNDHSRVYLMSLGK 1445

RESULT 11
Q9R0S4
ID Q9R0S4 PRELIMINARY; PRT; 1455 AA.
AC Q9R0S4,
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NIK-related kinase.
GN NRK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Kanai-Azuma M., Kanai Y., Okamoto M., Hayaashi Y., Yonekawa H.,
RA Yazaki K.,
RT "NrK: a murine X-linked NIK(Nck-interacting kinase)-related kinase
RL gene expressed in skeletal muscle.";
RL Mech. Dev. 89:157-161(1999).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB020741; BAA84943.1; -.
DR MGD; MG1:1351326; NrK.
DR GO; GO:0007256; P:activation of JUNKK; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
```



```
KW Hypothetical protein.
SQ SEQUENCE 292 AA; 33574 MW; 768168363DDACC2C CRC64;

Query Match 22.4%; Score 1551; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 1e-81;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1021 MLLDRSGGKVGGLIGRRRFQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLILHNDPE 1080
DB 1 MLLDRSGGKVGGLIGRRRFQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLILHNDPE 60

QY 1081 VEKKGQWTVGDMEGCHYRVVYKRIKFLVIALKSSVEVYAWAPKPYHKFMKAFKSPADL 1140
DB 61 VEKKGQWTVGDMEGCHYRVVYKRIKFLVIALKSSVEVYAWAPKPYHKFMKAFKSPADL 120

QY 1141 PHRPLLVDLTVEEGORLKYVYSSAGFAHVDVDSGNSYDIYIPVHIQSOITPHAIIFLPN 1200
DB 121 PHRPLLVDLTVEEGORLKYVYSSAGFAHVDVDSGNSYDIYIPVHIQSOITPHAIIFLPN 180

QY 1201 TDGEMMLLCYEDEGVYVNTYGRILIKDVVLQWEMPTSVAYICSNQIMGWGEKAIEIRSV 1260
DB 181 TDGEMMLLCYEDEGVYVNTYGRILIKDVVLQWEMPTSVAYICSNQIMGWGEKAIEIRSV 240

QY 1261 TGHLDGVFMHKAQRLKFLCERNDKVFASVRSVSGSSQVYFMTLNRNCIMNW 1312
DB 241 TGHLDGVFMHKAQRLKFLCERNDKVFASVRSVSGSSQVYFMTLNRNCIMNW 292

RESULT 13
QYUAN7 PRELIMINARY; PRT; 792 AA.
AC QYUAN7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Alternatively spliced serine/threonine protein kinase MIG-15
DE (Fragment).
GN MIG-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Pelodexinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Zhu X., Acharya P., Hedgecock E.;
RT "MIG-15, a NIK ortholog of the STR20 family of serine/threonine
RT protein kinases, is involved in cell migration and signal transduction
RT in C. elegans.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087132; AADI4594.1; -.
DR PIR; T43630; T43630.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
DR InterPro; IPR001180; Citron.
DR Pfam; PF00780; CNH; 1.
DR SMART; SM00036; CNH; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON TER 1
SQ SEQUENCE 792 AA; 87981 MW; EEE89627662742BE CRC64;

Query Match 21.0%; Score 1458; DB 5; Length 792;
Best Local Similarity 33.9%; Pred. No. 8.9e-76;
Matches 364; Conservative 127; Mismatches 242; Indels 340; Gaps 32;

QY 298 VRIQLKHIDRSRKRGKEETEYYSGBEDDHSHEGEPSSIMN-----VPGEST 350
DB 2 IRRHSKEHIDRNR--RVKDDADYSGSEDEDPSPNNRGPFGSMGRDRDSSSSMIPMDNT 59

QY 351 LRREFLRLQENKNSKALKQO-QOLQQOQRDPPEAHIKHLLHORRIEBOKE----- 403
DB 60 LRKGFQKLOESSRGFAEPGAQQLRLPOOPAP-----FOYQOSRYVEPRRESSEVKL 113
```

```
QY 404 ---ERRVVEOQRERERQRKLOEKEQORRLLED-----MQALRREERROAEREQEYKR 453
DB 114 RAVSSRGAADGPRHSPASPRPVSHQSRPOOSHAPAPHLADLANYEKRRESERE----- 168
QY 454 KOLEEQSERLORLOQOEHAHYLKSLOQQOQQOQLKQOQQOQLPGDRKPLHYHGRGNWP 513
DB 169 -----ERRERERQAHAM----- 181
QY 514 ADKPAWAREVEERTMNKQNSPLAKSKPGSTGPEPPIQASPPGPGPLSQTPPMQRP-- 571
DB 182 -----PIARVSASVPAP-QQSRKMSPELL 204
QY 572 ---VEPOEGPHKSLQDQPTRNLAAFPASHDPPALPAPTATPSARGAVIRONSDPTSGP 628
DB 205 ITHVKPED-----LDVLASELSKMGHH-----NGRSRES-- 234
QY 629 GPSNPNAWVRPDNEAPPKVPQRTSSATATLNTSCAGSRPAQAVRAPRPRNSAWQIY-- 686
DB 235 --SNSPP-----PPAPPREASISSITDTIDVG-----ELDNGADAEDDLKD 275
QY 687 LQRAAERTGPKPGPPAQPAPGPPNASSN--PDLRRSDPWERSDSVLPASHGHUPOAGSL 744
DB 276 IMMNGE--GTLRGNKPL--PPTPTDGTNTLVSDVRNGNG-----NSGH-----GAY 319
QY 745 ERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPPGPADEFVLLKERTLDEAPRPKAMDYS 804
DB 320 KGKKI-----PEIRPGIISLDD----- 337
QY 805 SSSEVESEDEDEEGEGGPAEGSRDTPGGRSDGDTDSVTMVVHDVEITGTPPYGGG 864
DB 338 -----DSDSNDEG----- 346
QY 865 TMVQRTPEERNLHADSNYTNLPDVQPSHSPSTNSKQSPSKDGDYQSRGLVVK 924
DB 347 ----NEPLMFKPINASSRG--ALPDLPLKSPQLRRQINDOTRQMSDDRADEQPNGFON 399
QY 925 APGKSSF-----TMFVDLGIYQPGSGSDIPITALVGGEGT-----RLDQLQ 966
DB 400 SDRSSSIQHSFNSRDREKSFVG---YFGGGAG-----AGGTVNRPGRPDQINQVQ 447
QY 967 YDVRKGSVNVNPTN--TRAHSETPEIRKYKRNSEITLCAALWGNLVLGTENGLMLLD 1024
DB 448 -----VWTPNSNGTPAENDAPEIRKYKKSGBILCAALWGNLVLITDSGLMLLD 499
QY 1025 RSGQKVGYLIGRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLILHND-----P 1079
DB 500 RSGQKVGYPILISRRRFDQMTVLEQNILATISGRKRIRVYVYLSWLRKQLIRTEGAGSAN 559
QY 1080 EVEKKQGWTVGDMEGCHYRVVYKRIKFLVIALKSSVEVYAWAPKPYHKFMKAFKSPAD 1139
DB 560 TTEKRWGNVNVGDIQGAHFKIVYERIKFLVWGLESSIEIYAWAPKPYHKFMKFSFGS 619
QY 1140 LPHRPLLVDLTVEEGORLKYVYSSAGFAHVDVDSGNSYDIYIPVHIQSOITPHAIIFLP 1199
DB 620 LSHVPLVDLTVEEDNARKLVLYGSTGGFHAIDLDSAAVYDIYTPAQSQTTTTHCIVVLP 679
QY 1200 NTDCHEMLLCYEDEGVYVNTYGRILIKDVVLQWEMPTSVAYICSNQIMGWGEKAIEIRSV 1259
DB 680 NSNCQMLLLCYDNEGVYVNTYGRMTKVVVLQWEMPSSVAYISTGQIMGWGKAIEIRSV 739
QY 1260 ETGHLDGVMHKAQRLKFLCERNDKVFASVRSVSGSSQVYFMTLNRNCIMNW 1312
DB 740 DTGHLDGVMHKAQRLKFLCERNDKVFSSAKGGSCQIYFMTLNRKPLTNW 792

RESULT 14
QYX4C4 PRELIMINARY; PRT; 728 AA.
AC QYX4C4;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
```


DE Alternately spliced serine/threonine protein kinase MIG-15
DN (fragment).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN SEQUENCE FROM N.A.
RP STRAIN=N2;
RC Zhu X., Acharya P., Hedgecock E.,
RA "MIG-15, a NIK ortholog of the STR20 family of serine/threonine
RT protein kinases, is involved in cell migration and signal transduction
RT in C. elegans."
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF087133; AAD14595.1; .
DR PIR; T43632; T43632.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
DR InterPro; IPR001180; Citron.
DR Pfam; PF00780; CNH; 1.
DR SMART; SM00036; CNH; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 728 AA; 80590 MW; 03FD87E04B839582 CRC64;

Query Match 19.9%; Score 1377; DB 51; Length 728;
Best Local Similarity 42.1%; Pred. No. 3.7e-71;
Matches 326; Conservative 94; Mismatches 204; Indels 150; Gaps 25;

QY 628 PGSP-----NPPAWVRDNEAPPKVPQRTSIALNTSGAGSRPAQAVRAPRNSNAWQ 684
DB 16 PAPAPFQYQSRVYPERES-SEVKRAVS-----SRGAADGRHSASRPPVPVSH--- 66

QY 685 IYLQRAERGTKPPGPPAPGPPNAS--SNPDLRSDPGWERSDVLDPASHGLPOAG 742
DB 67 ---QR-----SPOQSHPAAPHLADLANYEKRERSERERERERQAHHA-MPIA- 111

QY 743 SLERNRVGAS-----SKLDSFVPLSPGNKAKPDD-----HRSRGRPADPVLKERTL 790
DB 112 -----RVSAVSAPQQRKMSFLLI--THVKPEDLDVLASLSKVGHHNGRSREEMS 164

QY 791 DEAPRPKKAMDYSSSEVESE-----DDEE-----GEG-----GPAEGSRDTPGG 834
DB 165 PPPAPPAPPEASISSITDIDVGLDNGADAEDLDKIMNGEGTGLRGNKPLPPTP-- 222

QY 835 RSDGTDSTVSVVVDVVEITGTQPPYGGCTVWVORTPEERNLLHAD----- 882
DB 223 -----TDGENTLVSDVRRNGNS---GHGAYKGGKPIEIRPGIISLDDDDSDSDNEEGN 274

QY 883 -----SNGYTNLPDVQPSHSPTENSKSQSPSPKSGSD-----YQSRGLV 923
DB 275 EPLMPKPINASSRGALPDLIPKSPQLRRQINDQTRQMSDDRADESLSFGSYQPNQFG 334

QY 924 KAPGKSSP-----TMFVDLGIYQGGSGDSIPITALVGGEGT-----RLDQL 965
DB 335 NSDSRSSIQHSFNSDRKESFVG---YFGGGAG-----AGGTVNRRPGRPDINQV 382

QY 966 QYDVKRGSVVNVNPTN--TRAHSETPEIRKYKRRNSIILCAALGVNLLVGTENGLMLL 1023
DB 383 Q-----VNVTPNSNGTPAENDAPEIRKYKRFSGEILCAALGVNLLIGTDSGLMLL 434

QY 1024 DRSGQGVYGLIRRRFOQMDVLEGNLLITISGKRNKLRVYLSWLNRNKLHND----- 1078
DB 435 DRSGQGVYPLISRRRFDQMTVLEQNILATISGKRRIRVYLSWLNRNKLIRTEGASGA 494

QY 1079 PEVEKKQGGTVDMEGGCHYRVVYKRIKFLIVIALKSVEYAWAPKPYHKFAPKFSFA 1138
DB 495 NTTEKRNQVNVYDGLQGAHFVYRIKFLVWGLESSIEIYAWAPKPYHKFMSFKSFG 554

QY 1139 DLPHRLVLDLTVESQRLKYVSGAGFAVDVDSGNSYDIYIPVHQSQITPAHIFL 1198
DB 1198

DB 555 SLSHVPLIVLTVEDNARKLVYSGTGGFHAIDLDSAAVVDIYTPAOSGQTTTTHCIVVL 614
QY 1199 PNTDCEMLLCYEDGVVNTYIGRIIKDVVLQWCEMPTSVAYICNSNIMWGEXKAIEIRS 1258
DB 615 PNSNGMLLLCYDNEGVVNTYIGRTKVVVLQWCEMPTSVAYIYIQTQIMGNKKAIEIRS 674
QY 1259 VETGHLGDFVFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLNRNCIMNW 1312
DB 675 VDTGHLGDFVFMHKAQRLKFLCERNDKVFFSSAKGGSCQIYFMTLNKPGLTNW 728

RESULT 15
Q8C9S9
ID Q8C9S9 PRELIMINARY; PRT; 334 AA.
AC Q8C9S9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NIK related kinase (fragment).
GN NRK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RL EMBL; AK041377; BAC30923.1; .
DR MGI; MGI:1351326; Nrk.
DR GO; GO:0007256; P:activation of JUNK; IDA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase_1
DR PRINTS; PR00109; TRYKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
FT NON_TER 334
SQ SEQUENCE 334 AA; 37974 MW; DA40C1886FE503FF CRC64;

Query Match 14.3%; Score 993; DB 11; Length 334;
Best Local Similarity 56.1%; Pred. No. 1.9e-49;
Matches 185; Conservative 49; Mismatches 72; Indels 24; Gaps 1;

QY 1 MGDPAPARSLDDILSALRDPAGIFELVVGNGTYQGVYKGRHVKTGQLAAIKVMDVTE 60
DB 1 MAGPGSWRDXEVTDLQGLPDTGTGIFSLDKAIGLTGRIPLGIHEKTSGLVAVKVMARK 60

QY 61 -----DEEEIKQEIINMLKKYSHHNIATYYGAFIKKSPPG 96
DB 61 TPLPEIGRRVVRVVKYQKSVGRWYSDEEDLRLTELNLRLKYSFHKNIYTFYGAFFKLNP 120

QY 97 NDDQLWLVMEFCGAGSVTDLVQNTKGNALKEDCTAYICRETLGLAHHLAHKVIHRDIKG 156
DB 121 HQHQLWVWMLCAAGSVTDVVRNRQSLKEDWTAYICREILOGLAHHLAHKVIHRDIKG 180

QY 157 QNVLLITENAEVKLVDPGVSAQLDRTVGRNRTFTGTYYMAPEVIACDENPDATYDYSDI 216
DB 181 QNVLLTHDAEVKIVDFGVSAQVSRNTRGRNSFTGTYYMAPEVIHCDPDCSYDYSVDV 240

QY 217 WSLGITAIEMAEAGAPPLCDMHPRALFLIPRNPPLRKSKKSKKFDIDTCLIKTYLS 276
DB 276

Db 241 WSVGITAEMAGAPPLCKLQPLEALCVILREAPKVKSSGWSRKQNFQNMENCMIKNLF 300

Qy 277 RPTEQLLKFPFIRDQPTERQVRIQLKDH 306

Db 301 RPTSGNMLLHPFVHDIKNERRVVESLTKHL 330

Search completed: August 28, 2004, 01:00:27
Job time : 174 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:52:04 ; Search time 53 Seconds
(without alignments)
2381.194 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAAPSLDDIDLSALRD.....SGGSQVYFMTLNRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3951.5	57.0	1233	2 T30989	serine/threonine p
2	3008.5	43.4	1102	2 J0C316	probable protein k
3	2647.5	38.2	1075	2 T27623	hypothetical prote
4	2646	38.2	1080	2 T27622	hypothetical prote
5	2403.5	34.7	1027	2 T46481	hypothetical prote
6	1458	21.0	792	2 T43630	serine/threonine p
7	1377	19.9	728	2 T43632	serine/threonine p
8	710	10.2	836	2 B95716	probable serine/th
9	698	10.1	653	2 T34356	hypothetical prote
10	688.5	9.9	1233	2 T14157	serine/threonine p
11	678.5	9.8	426	2 S71886	Ste20-like protein
12	670	9.7	1231	2 T18532	serine/threonine pr
13	667	9.6	1080	2 S48944	hypothetical prote
14	658	9.5	1001	2 T17365	serine/threonine p
15	653.5	9.4	819	2 A53714	protein kinase (5C
16	651.5	9.4	1206	2 T34021	protein kinase SK2
17	640.5	9.2	829	2 T29372	hypothetical prote
18	635.5	9.2	690	2 C96572	protein F12M16.4 [
19	616.5	8.9	1135	1 A29813	132K ninaC protein
20	616.5	8.9	1501	1 B29813	174K ninaC protein
21	588.5	8.5	1228	2 T19897	hypothetical prote
22	582.5	8.4	1014	2 T31109	myosin III - Atlan
23	552.	8.0	471	2 T39232	probable serine th
24	551	8.0	982	2 T18576	serine-threonine k
25	548	7.9	490	2 S47946	protein kinase hom
26	535	7.7	312	2 T39525	serine/threonine p
27	533.5	7.7	658	2 T39500	serine/threonine-s
28	505	7.3	652	2 T39722	serine/threonine p
29	504.5	7.3	658	2 S60170	protein kinase Pak

30	501.5	7.2	561	2 T51417	protein kinase-lik
31	498.5	7.2	544	2 A57597	beta-p21-activated
32	497	7.2	710	2 T13458	hypothetical prote
33	496	7.2	544	2 S40482	serine/threonine-s
34	496	7.2	545	2 G01773	p21-activated prot
35	495.5	7.2	842	2 S60402	protein kinase CIA
36	493.5	7.1	544	2 I49376	p21 activated kina
37	489	7.1	553	2 T01479	hypothetical prote
38	488	7.0	589	2 T38086	serine/threonine-p
39	481.5	6.9	525	2 S58682	protein kinase, p2
40	478.5	6.9	622	2 T15467	hypothetical prote
41	478.5	6.9	655	2 S51884	probable protein k
42	474.5	6.8	378	2 T26684	hypothetical prote
43	463	6.7	693	2 B85112	hypothetical prote
44	459	6.6	939	2 S28394	probable serine/th
45	455	6.6	607	2 T01904	hypothetical prote

ALIGNMENTS

RESULT 1

T30989
serine/threonine protein kinase NIK - mouse
N/Alternate names: Nck interacting kinase
C/Species: Mus musculus (house mouse)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C/Accession: T30989
R/Su, Y.C.; Han, J.; Xu, S.; Cobb, M.; Skolnik, E.Y.
EMBO J. 16, 1279-1290, 1997
A/Title: NIK is a new Ste20-related kinase that binds NCK and MEKK1 and activates the S
A/Reference number: Z20954; MUID:97280817; PMID:9135144
A/Accession: T30989
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1233 <SU>
A/Cross-references: EMBL:U08984; NID:g1872545; PID:g1872546; PIDN:AAC53165.1
C/Keywords: protein kinase

Query Match 57.0%; Score 3951.5; DB 2; Length 1233;
Best Local Similarity 62.3%; Pred. No. 1.8e-105;
Matches 846; Conservative 112; Mismatches 228; Indels 173; Gaps 32;

Qy	1	MGDPAPARSLDDIDLSALRDPA	GIFELVEVVGNGTYGQVYGRHVKTQ	LAALKMVDYTE	60
Db	1	MANDSPAKSLVDIDLSSLRDP	AGIFELVEVVGNGTYGQVYGRHVKT	-VTAAIKMVDYTE	59
Qy	61	DEEEIKQEIINLKKYSHHRNI	ATYGAFAFKKSPGNDQDLWVMEFCG	AGSVTDLVKNT	120
Db	60	DEEEITLLEINLKKYSHHRNI	ATYGAFAFKKSPGDDQDLWVMEFCG	AGSITDLVKNT	119
Qy	121	KGNALKEDCIAYICREILRL	GLAHLHAHKVIRHDIKGNVLLT	ENAEVKLVDFGVSQAQLDR	180
Db	120	KGNLIKEDWIAVISREILRL	GLAHLHIHVIRHDIKGNVLLT	ENAEVKLVDFGVSQAQLDR	179
Qy	181	TVGRNRTFPGPYMAPEVIA	CDENPDATYRSDIWSLGITAE	MAGAPPLCDMHPMR	240
Db	180	TVGRNRTFPGPYMAPEVIA	CDENPDATYRSDLSMCGITAE	MAGGPPPLCDMHPMR	239
Qy	241	ALFLIPRPPRLKSKWKKFI	DDTCLTKLYLSRPPTQLK	FPPIRQPTQVR	300
Db	240	ALFLIPRPPRLKSKWKKFF	SFIEGLVKNYMRPSTQLK	HPPIRQPNRQVR	299
Qy	301	QLKDHIDRSRKRGKEBTE	YEYSGSBEEDDSHGE-EGEP	SSIMNVPGESTLRREFLR	359
Db	300	QLKDHIDTRKRGKEBTE	YEYSGSEEEVEPQEGEP	SSIVNVPGESTLRRLRFL	359
Qy	360	QENKSNFALKQOQOQO	QOQOQDPPEAHIKHLLHQR	ORRIEIQKEGERRRVEQQR	419
Db	360	QENKSEALRQQLLOEQ	QLREQEYKRLLAERQKIEQ	QKEQRRLLEEQRREAR	419
Qy	420	KLQKEQ-----QR	LEDNQALRR-----EEER	QRAERQEVKRLKOL-EEQ	463
		:	:	:	

Qy	638	VRPDNEAPKVPKPORTSSIIATALNTSGAGSRPAQAVRAPRPSNSAWQIY--LQRRABRG	695
Db	528	-----PPAPPREASISITDITDVG-----ELONGADAEMDLDKDIMNGE-GT	571
Qy	696	PKPGPPAOPPGPPNASSN--PDLRSDPCWERSDSVLPAASHGHLPAQAGSLERNRVGASS	753
Db	572	LRFGNKPL-PPPTDGTENTLVSVRNNG-----NSGH-----GAYGKKI----	612
Qy	754	KLDSSFVLSPGNKAKDPDHRSRPRPADFVLLKERTLDEAPRPKKAMDYSSSSSEEVSS	813
Db	613	-----FEIRPGIISLDD-----	627
Qy	814	EDDEREGEGPARGSRDTPGRRSDGTDTSVSTVMVHDVBEITQTQPPYGGTVMVORTPE	873
Db	628	DSDNREG-----	638
Qy	874	EERNLLHADNSGYTNLPDVVPVQSHSTENSKGQSPSKDGSQYQSRGLVKAPGKSPF--	931
Db	639	MFKPINASSRG--ALPDLPLFKSPQLRRINDQTRQMSDDRADEQNGFONSDRSISQH	696
Qy	932	-----TMFVDLGIYQPGSGDSIPITALVGEGT-----RLDQLQYDVRKGSVV	975
Db	697	SFSDNRDREKSFVG--YFGGAG-----AGGTVNRPGRPDQINQVQ-----V	736
Qy	976	NVNPTN--TRAHSETEIRYKRFNSILCAALGWNLVLGVTEGLMLLDRSGQGVYK	1033
Db	737	NVTPNSNGTPAENDAPFIRKYKKFKSGEILCAALGWNLIIGTDSLGLMLLDRSGQGVYK	796
Qy	1034	LIGRRFPQMDVLEGNLLITTSKRNKLRVYVLSWLRNKLHND-----PEVEKQCGWT	1088
Db	797	LISRRFPDQMTVLEGQNILATISGRKRRIRVYVLSWLRQKILRTEGAGSANTTTEKRGWV	856
Qy	1089	TVGDMGCGHYR.VKYERIKFLVIALKSSVEVYAWAPKPVHKFMFKSPADLPRLPLVD	1148
Db	857	NUGDLQALHFKLIVRIKFLVVGLESSIEITYAWAPKYHKFMFKSPGSLSHVPLVD	916
Qy	1149	LTYEEOQLKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPNTDGMEMLL	1208
Db	917	LTVEDNARLKVLYGSTGGFHAIDLDSAAVTDIYTPAQSGQTTTTPHCIVVLPNSNGMQLLL	976
Qy	1209	CYDEEGVYNTYGRITIKDVLQWGENPTSAVYICSNQIMGWGEKATEIRSVETGHLGVF	1268
Db	977	CYDNEGYYNTYGRMTKNVYLQWGENPSSVAYISTQIMGWGNKATEIRSVDTGHLGVF	1036
Qy	1269	MHKRAQLKFLCBERNDKVFFASVRSGSSGVYFMTLNRNCIMNW	1312
Db	1037	MHKKAQLKFLCBERNDKVFFSNAKGGSGCQIYFMTLNLKPLTNW	1080

RESULT 5

T46481

hypothetical protein DKFPz434A025.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C;Accession: T46481

R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z21035

A;Accession: T46481

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1027 <AAA>

A;Cross-references: EMBL:AL137755

A;Experimental source: adult testis; clone DKFPz434A025

C;Genetics:

A;Note: DKFPz434A025.1

Qy	262	FIDFIDTCLIKTYLSRPPTQQLKFPFIRDQPTROVRIQLKDHDNRKRGKEETEY	321
----	-----	--	-----

Query Match 34.7%; Score 2403.5; DB 2; Length 1027;

Best Local Similarity 51.1%; Pred. No. 1e-61;

Matches 560; Conservative 108; Mismatches 237; Indels 191; Gaps 34;

Db 1 PPSPIEGCLVKNYMPSTEQLLKHPFIIDQPNRQVRIQLKDHIDTRKKRGEKDETEY 60

Qy 322 EYSGSEEDDSHGE-EGPSSIMVYSGSTLRLREFLRLOQENKNSSEALKKQOQLOOQOQ 380

Db 61 EYSGSEEBEEVEPEGEPSIVNVYSGSTLRLREFLRLOQENKNSSEALKRQOQLOOQO 120

Qy 381 RDPKAIHKLHQRORRIEEOKEERRRVEEQORRERQKLOEKEO-----QRLRDM 433

Db 121 REQEYKQLLAERQKLEQOKEQRRLLEEORRERARQOERREORREOEKRLBEL 180

Qy 434 QALRE-----EEREOAREOEYKQKLEQO----- 459

Db 181 ERRKEEERRRABEKEKREVEQEYIRQLLEEQRHLEVLQOQLLOEQMLLECRWEM 240

Qy 460 ---ROSELRQLOQEHAYLSLQO-----QOQOQLOKQOQOQLLPGDR--KPLYHYG 508

Db 241 EHRQAEERLQLOQEQAYLLSLQHDHRRPQHSQOQPPPPQOQSKPSFAPEPKAHY- 299

Qy 509 RGMNPADKPAWAEVEERTMKNQONSPLAKSPGSTGPEPPIQASFGPGLSQT--P 566

Db 300 ---EPADR---AREVEDRFR-KTNHSSPEAOSKQGRVLEPPVPVPSRSFSFNGNSSEVHP 352

Qy 567 PMQREVEPO-EGPH-KSLQDQPTRNLAAFPASH---DPDPA-----IPATATPSARGAV 616

Db 353 ALQRAEAEQVQWHLASLKN-----NVSPVSRSHSFSDSPKFAHHLASQDPCPPSRSEV 408

Qy 617 IRONSDDPTSEGGPSPNPAPWRPD-NEAPPKVPQORTSSITATANT-----SGAGGSR 668

Db 409 LSQSDSKSEAPDPTOK--AWSRSDSDVPPRPVVRTTSRSPVLSRRDSPSLQSGSQNSQ 466

Qy 669 PAQAVRPRNSAWQIYLORAEGRTPKPPGPAQPPGPNASSNPDLRSDPCWERSD 728

Db 467 AGQ-----RNSTSSIEPRLLWERVEKLVPKPGS--GSSSGSNSGSPQG---SHPG----- 512

Qy 729 SVLPASHGHLPOAGSLERNVGCSSKLDSSPVLSPCNKA-KPDDHRS--RQGRPADFY-- 783

Db 513 -----SOSGGEFRFRVSSSKSEGPSORLENVAKPEDKKEVFRPLKPAGEVDL 562

Qy 784 --LLKERTLEAPRPKPKAMDYSSSEV-----ESSEDEEAGEGPAEGRDTPGGR-- 835

Db 563 TALAKELRAVEDVRPHKVTDYSSSESGTTDEEDDVEQEGADESTSGPEDTRAASSL 622

Qy 836 --SDGDTSVSTWVHDVBEITGTOPPYGGGTWVQRTPEBERNLLHADSNGYTNLDPV 893

Db 623 NLSNGETESVKTMIVHDDVSEPAWTPSKBGLTVRQT----- 660

Qy 894 QPSHSTPENSKGSPSPKSGSDYQSRGLVAPKGSFTMFVDLGIYOPGGSGDIPITA 953

Db 661 -----QSASST-----LQHKSSSSFTFPDPRLLQISPSGT-TWIS 697

Qy 954 LVG--GEGTRLDQLYD-VRKGSVVNVNPTNTRAHSETPEIRKYKRFNSEILCAALMGV 1010

Db 698 VVGPSCDGMRPEAIRQDPTKGSVVNVNPTNTRPQSDTPEIRKYKRFNSEILCAALMGV 757

Qy 1011 NLLVGTENGLMLLDRSGQGVYLLGRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWL 1070

Db 758 NLLVGTESGLMLLDRSGQGVYLLNRRRFOQMDVLEGLNLLVITISGKKDLRYVYLSWL 817

Qy 1071 RNKILHNDPEVEKKQWTTVGDMEGCGHVRVVKYERIKFLVIALKSSVEVYAWAPKPYHK 1130

Db 818 RNKILHNDPEVEKKQWTTVGDMEGCGHVRVVKYERIKFLVIALKSSVEVYAWAPKPYHK 877

Qy 1131 FMAFKSFADLPHRPLLDVLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSI 1190

Db 878 FMAFKSFGLVHKPLLDVLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSI 937

Qy 1191 TPFAIIFLNTDGMELLCYEDEGVYNTYGRHIIKDVVLQWG-EMPTSVAY----- 1240

Db 938 KPFAIILPNTDGMELLCYEDEGVYNTYGRHIIKDVVLQWG-EMPTSVAY----- 997

Qy 1241 -----ICSNQIMGW 1249

Db 998 REGHRDPICGNWSLGM 1013

RESULT 6

T43630

serine/threonine protein kinase homolog - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C:Accession: T43630

R:Zhu, X.; Acharya, P.; Hedgecock, E.

submitted to the EMBL Data Library, August 1998

A:Description: MIG-15, a NIK ortholog of the STE20 family of serine/threonine protein ki

A:Reference number: 222584

A:Accession: T43630

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-792 <ZHU>

A:Cross-references: EMBL:AF087132; PIDN:AAD14594.1

C:Genetics:

A:Map position: X

A:Note: mig-15

Query Match 21.0%; Score 1458; DB 2; Length 792;

Best Local Similarity 33.9%; Pred. No. 4.7e-35;

Matches 364; Conservative 127; Mismatches 242; Indels 340; Gaps 32;

Qy 298 VRIQLKHIDRSKRKGEKETEYEGSSEEDDSHGEGEPSSIMN-----VPGEST 350

Db 2 IRHSIKHEIDRNR--RVKDDADYEYSGSEDDPSNNRGPMSGIRDDSSSSMIPDMT 59

Qy 351 LRREFLRLOQENKNSSEALKKQ--QOQOQOQORDPEAHIKHLLHQRORRIEKE----- 403

Db 60 LRKGFQKLQSSRGFAEPGAQQLRLPQAPAP-----FQYQOSRVPRRESSEVKL 113

Qy 404 ---ERRRVEEQRRERQKLEKEQQRRLD-----MQALRRREERRRQAEERQYKR 453

Db 114 RAVSSRGADPRHSPASPRPVSHQKSPQSHPAAPHLADLANYEKKRRSRE----- 168

Qy 454 KQLEEQRSERLQLOQEHAYLSLQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 513

Db 169 -----ERRERQAHAM----- 181

Qy 514 ADKPAWAEVEERTMKNQONSPLAKSPGSTGPEPPIQASPGPPGLSOTPPMORP-- 571

Db 182 -----PIARVSASVPAP-QQSRKMSPELL 204

Qy 572 ---YEQEGPHKSLQDQPTRNLAAFPASHDPPAIPAPTATPSARGAVIRQNSDPTS 628

Db 205 ITHVKPED-----LDVLASELSKMGHH-----NGRSRE-- 234

Qy 629 GPSNPAPWVRPDNEAPKVPQORTSSITATNTSGAGSRPAQAVRPRNSAWQIY-- 686

Db 235 --SNMPP-----PPAPPPREASISSITDIDVG-----ELDNGADAEMDDLDK 275

Qy 687 LQRAERGTKPPGPPAPQPPPNASN--PDLRRSDPGWERSDVLPAASHGLPQAGSL 744

Db 276 IMMNGE--GTLGPNKPL--PPTPTDGTWLVSDVRNGNG-----NSGH-----GAY 319

Qy 745 ERNRVGAASKLDSSPVLSPGNKAKPDHRRSRGPADFVLLKERTLDEAPRPPKAMDYS 804

Db 320 KGKKI-----PEIRPGIISLDD----- 337

Qy 805 SSSSEVESSEDEEGEGGPAEGRDTPGSRSDGSDTSVSTWVVDVEITGTQPPYGGG 864

Db 338 -----DSDSDNEEG----- 346

Qy 865 TMVQRTPEBERNLLHADSNGYTNLPDVQPSHSPTENSKGQSPSKDGGDYQSRGLVK 924

Db 347 -----NPLMKPINASSRG--ALPDLPLKSPQLRQINDQTRQMSDDRADEQNGFN 399

Qy 925 APGKSSP-----TMFVDLGIYOPGGSGDSIPITALVGGEST-----RLDQLQ 966

Db 400 SDRSSSIQHSFNRDRKESFVG---YFGGAG-----AGGCTVNRPGRPQDINQVQ 447

Db 416 EVI---QEN---RYDKVDVWALGVSAIEMAGELPPRSSVHPNRVLFMISIEPAPMLEDK 470
Qy 257 KWSKXFPIDFIDCLIKTYLSRPTTQLKFPPIRDOPTQVRIQLKDHIDRSRKRKG-- 314
Db 471 KWSLVFHDFAKCLTKPEPLRPTAAEMLKHKEVERCKTGASA---MSPKIEKSRQIRATM 527
Qy 315 -----EKEETEEYSGSEEDDSDHGESEGESESSIMNVPGESTLREFLR----- 357
Db 528 ALQAQSVVAPSLSDTSLGPKSSELGITVPSPKPPONSTEAPLTSTLNQRHITGNVTLAG 587
Qy 358 -----LQOENKNSSEALKQOQQLQOQOQORDP 383
Db 588 EGGDFGTWIVHGEDETESDRSQLVREKSSSQFEGVP 627

RESULT 9

T34356
Hypothetical protein T19A5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34356
R:Bradshaw, H.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid T19A5.
A:Reference number: Z21512
A:Accession: T34356
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-653 <BRA>
A:Cross-references: EMBL:U53153; PIDN:AAC69038.1; GSPDB:GN00023; CESP:T19A5.2
A:Experimental source: strain Bristol N2; clone T19A5
A:Genetics:
A:Gene: CESP:T19A5.2
A:Map position: 5
A:Introns: 26/3; 101/3; 157/1; 209/3; 272/3; 305/2; 373/3; 635/3

Query Match 10.1%; Score 698; DB 2; Length 653;
Best Local Similarity 43.7%; Pred. No. 1.3e-13;
Matches 150; Conservative 64; Mismatches 97; Indels 32; Gaps 9;
Qy 3 DPAPARLDD-----IDLALRDAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVM 56
Db 7 DELPQADDDSMKWRRIYQKL-DPEVFTKQERIGRSGFGEVYKGINRTGRVVAIKII 65
Qy 57 DV--FEDEEEIKQBINMLKKTSHRNATYYGAFIKKSPGNDQMLVMVEFCGAGSVT 114
Db 66 DLEQAEDEIEDIQIQLVLSQ-CDSQYVTKYSGFLKGS-----KLMIIMEYLGGSAL 118
Qy 115 DLVKNKGNALKEDECIAYICREILGLAHLHAHKVHRIKQNVLLTENAERVKLVDFGV 174
Db 119 DL---TKSGKLDESHIAVILREILKGLVYLHSEKIHRIKAAANVLVSEHGDKVADFV 175
Qy 175 SAQLDRTVGRRTFTGTPTWMAPEVIACDENPDATDYRSDIWSIGITAIEMAGAPLIC 234
Db 176 AGQLTETVKRITFTVGSPPWMAPELI-----KQSSDYDKADISLIGITAIELANGEPHS 230
Qy 235 DNHPRALFLIPNPPRLKSKWKKFIDFIDTCLIKTYLSRPTTQLKFPPIRDOPT 294
Db 231 DLHPNRVFLIPNPPVVLQSQWQSKPFKEFVEMCLNKDPENRPSASTLLKHQFIK-RAK 289
Qy 295 ERQVRIQLKDHIDRSRKRKEETEYYSGSEEDDSDHGE 337
Db 290 KNSILVDLIERAAEYRLRTGVS-----SDSLDESDSGGG 325

RESULT 10

T14157
serine/threonine protein kinase - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14157
R:Pyotowski, B.; Hicklin, D.J.; Kornhaber, G.; Dellaratta, D.V.; Witte, L.

submitted to the EMBL Data Library, December 1997
A:Reference number: Z17894
A:Accession: T14157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1233 <PVT>
A:Cross-references: EMBL:AF039574; NID:g2773155; PID:g2773156; PIDN:AAB96682.1

Query Match 9.9%; Score 688.5; DB 2; Length 1233;
Best Local Similarity 26.0%; Pred. No. 4.1e-13;
Matches 236; Conservative 145; Mismatches 333; Indels 193; Gaps 37;

Qy 20 DPAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMVDTEDEE-EETKQINMLKKTYS 78
Db 29 NPEFWELIGELGDAFGKVKYAKQKNTVLAAKVIDTKSEBEEUWVIDILAS-CD 87
Qy 79 HRNIATYYGAFIKKSPGNDQMLVMVEFCGAGSVTDLVKNKGNALKEDECIAYICREIL 138
Db 88 HPNIVKULDIFY-----YENNLWILIEFFCAGGAV-DAVMLELERPLTESQIVVCKQTL 140
Qy 139 RGLAHLHAHKVHRIKQNVLLTENAERVKLVDFGVSAQLDRTVGRRTFTGTPTWMAPE 198
Db 141 EALNYLDHNKIHRDLKAGNLTLDGDIKLADFGVSAKNTRTTIQRDPSFTGTPYMAPE 200
Qy 199 VIACDENPDATDYRSDIWSIGITAIEMAGAPLICDHPMRALFLIPNPPRL-KSKX 257
Db 201 VMCETSKDRPYDYKADVWSLIGITLIEWALEPHEHLMNMRVLLKIAKSEPTLAQPSK 260
Qy 258 WSKKFIDPIDCLIKTYLSRPTTQLKFPPIRDOPTQVRIQLKDHIDRSRKR----- 313
Db 261 WSSNFKDPLKCLEKNVDARWTTSQLQHPFVT-----VDSNKPVELIA 305
Qy 314 -GEKEETEYYSGSEEDDSDHGESEGESESSIMNVPGESTLREFL-LRLOQENKNSSEALKQ 371
Db 306 EAKAEVTEEVBDGKEDEEEEAENALP-----IPANKRASSDLSIASSEEDKLSQACIL 360
Qy 372 QOQLQOQOQORDEAHKH-LHQR-----QRRLEEOK-----BERRVEEQ-----QR 413
Db 361 ESVSERTEQSTSEDKFNKILNEKPTDGPKEKAVDEHASDVNLETGAELNDQTVGIHENG 420
Qy 414 RERQRLQKE--KEQORRLEDQALRREERERQAREQEYKRRKQLEEQSERLQRLQO 471
Db 421 REKKRPLENLPTDQDQTVDSVSEENNRVLTNTDCLKPEEDRNKE-----NQ 474
Qy 472 EHAYLKSLLQOQO-----QOQOQOQOQOQLLPGDKPLHYHGRGNPADPAMAREVEE 525
Db 475 ETLESKLITQSEINDTHITMDLVSQE-----TGEKEADFQAVDN 514
Qy 526 RTRNNKQNSPLAKSKPGSTGPEPPIQASPGPPGLSQTPPMORPVEPQSGPHKSLQD- 584
Db 515 EVGLTKSET-----QEKLGKDGTAQKVITSDRSSE--VGTDEALDDT 554
Qy 585 QPTRNLAAFPASHDPDAIPAPTATPSARGAVIRQNSDPTSEGP---GPSNPPPAWVRPD 641
Db 555 QKAELSKAASGSGDEAL-VPTQTLEAK-----PT-EGEAGAGAESEPPGGERVE 603
Qy 642 NEAP---PKVQRTSSITATNTSAGGSRP-AQAVRARPRSNASAWIYLQRRARERTPK 697
Db 604 DKQEQEQPAVCEAEGQLTSETTRATLEQPETDEVEQVSESNSIEE--LERLVVTGA-- 659
Qy 698 PPGPPAPQPPGPPN--ASSNPDLRRSDPCWE---RSDSVLPASHGHLPOAGSLERRNRVGN 752
Db 660 ----EARALGSEGAATAEVDLERKENAQKVFVKAESQAPA-----AS 698
Qy 753 SKLDSSPVLSPCNKAKPDHRSRPCRADFVLLKERTLDEAPRPKKKAMDYSSSEEVES 812
Db 699 QPSFPHVPLIPNINSETTENKEEMGA--LPKETI---LPP-----EPEHE 741
Qy 813 SEDDEEGEGGPAEGSRDTPGGRSDGDTD-SVSTMVHVDVEEITGTQPPYGGGTWVORT 871
Db 742 KGNDTSDGTGTVSE-----NSSGDLNLSISSFL-----SKAKDSGVSLOQET 783
Qy 872 PEERNL 878


```
QY 298 VRIQLKDHIDRSKRKRGEETEYE-----YSG-----SEEDDSHG----- 334
Db 287 -ETVLIDLITQTKDAVDRLNLDQYKMKLLFOEAHNGPVAQAEHEEDHGGRTGTV 345
QY 335 -----EEGPPSIMNVPGESTLRREFLRLOQEN--KSNSEA--LKQOQOL 375
Db 346 NSVGNQSPISMSISASSQSSVNSLPDASDDKSLDMWEGDHTVMSNSSVTHLKPEE- 404
QY 376 QOQQRDPENHI-----KULLHOR-----RIEQKEERRRVEEQ- 411
Db 405 NYQEEGDPRTRASAPQSPQVSRHKSHYRNREHFATIRTSALVTROMQEHQDSLRQOM 464
QY 412 ----QRREREQKLOEQEQRLE--DMOALRREE------RRAEREQ 449
Db 465 SGYKEMRQHOKQMTLENKLAENDEHRLDKDLETQRNFAEMEKLIKQHASMEK 524
QY 450 EYRKOLBEQROSERLQRLQOE--HAYLKS-----LQOQOQOQOLQ-----KQOQOOL 496
Db 525 EAKVMANBEKKPQHQIQAQKKELNSFLESQKREYKLRKEQLKELNENQSTPKKEKQEW 584
QY 497 LPQDRKPLYHYGRGNPNADKPAAWAVEERTMKNQNSPL 537
Db 585 LSKQENIQHF-----QAEENLRRRQYL 611

RESULT 15
AS3714
Protein kinase (EC 2.7.1.37) BL44 - human
N:Alternate names: GC kinase
C:Species: Homo sapiens (man)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Sep-1999
A:Accession: A53714
R:Katz, P.; Whalen, G.; Kehrl, J.H.
J. Biol. Chem. 269, 16802-16809, 1994
A>Title: Differential expression of a novel protein kinase in human B lymphocytes. Prefe
A:Reference number: A53714; MUID:9426900; PMID:7515885
A:Accession: A53714
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-819 <KAT>
A:Cross-references: GB:U07349; MID:g531819; PIDN:AAA20968.1; PID:g531820
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; phosphotransferase
F:13-272/Domain; protein kinase homology <KIN>
F:21-29/Region; protein kinase ATP-binding motif

Query Match 9.4%; Score 653.5; DB 2; Length 819;
Best Local Similarity 29.8%; Pred. No. 2.9e-12;
Matches 186; Conservative 84; Mismatches 181; Indels 173; Gaps 23;

QY 17 ALRDPAGIFELVVGNGTYGVYKGRHVKTQLAAIKVMDVTE--DEEEIKQEIINLKK 75
Db 7 SLQDRPRDFELLQVRGAGTYGVYKARDVTVTSELAAVKIVKLDPGDDISSIQQEITILRE 66
QY 76 YSHENIATYYGAFIKKSPGNDQDLWLMFPCGAGSVTDLVKNTKGNALKEDCIAYICR 135
Db 67 -CRHFNVAIYGSYL-----NDRLMICMEFCGGSLSQEIYHAT--GPLEERQIAYVCR 117
QY 136 EILRGLAHLHAHKVTHRDIKGNVLLTENAENVKLVDGVSQQLDRVTGRRNTFTGTPYWM 195
Db 118 ERLKGLHLHLSQKTHRDIKGANLLLTQGDVKLADFGVSGELTASVAKRRSFITGTPYWM 177
QY 196 APEVTACDENPDATDYDSIDWSLGLTAEIEMAEAGAPPLCDMHPMEALFIPN--PPRL 253
Db 178 APEVAARERK--GGYNEICDWMALGTAIELEGLOPFLHLHPMHALMLMSKSSFPQPKL 235
QY 254 KSK-KWSKKFIDFIDTCLIKTYLSRPTEQLIKFPFIRDQDPTQROVRIQLKDHIDRSRKX 312
Db 236 RDKTRWTQNFHFLKALTAKPKKTAELKLOHPTTQQ-LPRALLTQLLDKASDPHLG 294
QY 313 RGEKEETEYE--YSGSEEDDSHGEGEPSSIMNVPGESTLRREFLRLOQENKNSSEALKQ 371
```

```
Db 295 TPSPEDCELETYDMFPDTIHSRGQH-----PAERT----- 325
QY 372 QOQLQOQOQORDPFAHIKHLHQRORRIEEQKEERRRVEEQRREREQKLOEKEQORRELE 431
Db 326 PSEIQ-----FHQ-----VKFGAPRRKETDPLNEPWE-----E 353
QY 432 DMQALRREERRRQAEREQEVKRYKQLEEQROSERLQRLQOEHAHYLKSLLQOQOQOQLOKQ 491
Db 354 EWTLLGKEE-----LSGSLLQSVQEALEERSLTIR 383
QY 492 QOQOQLLPQDRKPLYHYGRGNPNADKPAAWAVEERTMKNQNSPLAKSKFGSGTGPPEPI 551
Db 384 SASEFQELD-----SPDD-----TMGTIKRAPFL-----GPLFTDPPA 416
QY 552 PQASGPPGPIQSOTPPMQRPVPEQEGPHKSLQDQPTRNLAAFPASHDDPDAIPAPTATPS 611
Db 417 EPLUSSPPGTL-----PP-----PPSGNNS-----PLLTAWAT-- 446
QY 612 ARGAVIRQNSDP-----TSEGGPQSP 632
Db 447 -----MKQREDPERSSCHGLPPTP 465
```

Search completed: August 28, 2004, 01:01:22
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:57:10 ; Search time 81 Seconds
(without alignments)
5095.952 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAARSLLDIDLSALRD.....SGSSQVYFMTLRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6929	100.0	1312	14	US-10-029-115-2
2	6660.5	96.1	1295	10	US-09-789-390-30
3	6660.5	96.1	1295	10	US-09-789-390-32
4	6660.5	96.1	1295	10	US-09-789-390-34
5	6660.5	96.1	1295	10	US-09-789-390-37
6	6660.5	96.1	1295	10	US-09-789-390-39
7	6659	95.1	1326	10	US-09-291-417-15
8	6646.5	95.9	1303	10	US-09-789-390-35
9	6646.5	95.9	1303	10	US-09-789-390-38
10	6628.5	95.7	1303	10	US-09-789-390-9
11	6622	95.6	1332	10	US-09-789-390-7
12	6603.5	95.3	1276	14	US-10-029-115-6
13	6311	91.1	1244	10	US-09-789-390-13
14	6304.5	91.0	1273	10	US-09-789-390-11
15	5422.5	78.3	1203	10	US-09-789-390-65

16	5412	78.1	1332	10	US-09-789-390-63	Sequence 63, Appl
17	5261	75.9	1244	10	US-09-789-390-70	Sequence 70, Appl
18	5261	75.9	1244	10	US-09-789-390-71	Sequence 71, Appl
19	5250.5	75.8	1273	10	US-09-789-390-67	Sequence 67, Appl
20	5250.5	75.8	1273	10	US-09-789-390-68	Sequence 68, Appl
21	4336	62.6	1360	9	US-09-871-916-2	Sequence 2, Appli
22	4336	62.6	1360	14	US-10-355-975-14	Sequence 14, Appl
23	4274	61.7	1268	15	US-10-353-690-122	Sequence 122, App
24	4195.5	60.5	1297	10	US-09-291-417-14	Sequence 14, Appl
25	4021.5	58.0	1239	10	US-09-291-417-13	Sequence 13, Appl
26	3977	57.4	1212	12	US-10-168-582-9	Sequence 9, Appli
27	3977	57.4	1212	14	US-10-247-671-157	Sequence 157, App
28	3951.5	57.0	1233	10	US-09-291-417-89	Sequence 89, Appl
29	3872.5	55.9	1165	12	US-10-211-462-89	Sequence 89, Appl
30	3872.5	55.9	1165	14	US-10-021-660-126	Sequence 126, App
31	3824	55.2	1175	9	US-09-771-161A-224	Sequence 224, App
32	3824	55.2	1175	9	US-09-771-161A-225	Sequence 225, App
33	3824	55.2	1175	9	US-09-771-161A-226	Sequence 226, App
34	3821	55.1	792	14	US-10-029-115-4	Sequence 4, Appli
35	2708.5	39.1	1109	10	US-09-291-417-88	Sequence 88, Appl
36	2403.5	34.7	1027	16	US-10-408-765A-1181	Sequence 1181, Ap
37	2280	32.9	510	9	US-09-925-297-596	Sequence 596, App
38	1862.5	26.9	1227	10	US-09-291-417-105	Sequence 105, App
39	1819.5	26.3	1581	12	US-10-415-011-16	Sequence 16, Appl
40	1797	25.9	515	9	US-09-771-161A-134	Sequence 134, App
41	1686	24.3	425	9	US-09-771-161A-133	Sequence 133, App
42	1676.5	24.2	468	9	US-09-771-161A-135	Sequence 135, App
43	1600.5	23.1	339	16	US-10-664-421-79	Sequence 79, Appl
44	1532	22.1	288	12	US-10-276-774-1920	Sequence 1920, Ap
45	1457	21.0	275	14	US-10-355-975-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-029-115-2
; Sequence 2, Application US/10029115
; Publication No. US20030077597A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Fu, Alan C
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: NO. US20030077597A1el Germlinal Center Kinase Cell Cycle Protiens,
; FILE REFERENCE: A-70229/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/029,115
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-115-2

Query Match	100.0%;	Score 6929;	DB 14;	Length 1312;
Best Local Similarity	100.0%;	Pred. No. 4.4e-311;		
Matches 1312;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	MGDPAPARSLLDIDLSALRD	PAGIFELVEVVGNGTYGQVYKGRHVKTGQLAAIKVMDVTE	60
Db	1	MGDPAPARSLLDIDLSALRD	PAGIFELVEVVGNGTYGQVYKGRHVKTGQLAAIKVMDVTE	60
Qy	61	DEEEIKQEIINMLKKYSHHRIATYYGAFIKKSPGNDQDLWLVMFCGAGSVTDLVKNT		120
Db	61	DEEEIKQEIINMLKKYSHHRIATYYGAFIKKSPGNDQDLWLVMFCGAGSVTDLVKNT		120
Qy	121	KGNAKEDCTIAYICREILRGLAHLHAHKVHRDIKQGNVLLTENAIVKLVDFGVSQALDR		180
Db	121	KGNAKEDCTIAYICREILRGLAHLHAHKVHRDIKQGNVLLTENAIVKLVDFGVSQALDR		180
Qy	181	TVGRNTFIGHTPYWMAVEVACDENPDATYDYSIDISLGTATEMAEGAPPLCDMHPMR		240

Db 481 QQQQQQLOKQQQQQLPGDRKPLHYHGRGMNPKADPAWAVEVERTRMKNQKNSPLAKS 540
Qy 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMORPVEPQEGPHK----- 580
Db 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMORPVEPQEGPHKSLVAHRVPLKPYAAPVPRSQ 600
Qy 581 SLQDQPTNLAAFPASHDPPAI PAPTATPSARGAVIRQNSDPTSEGPGSPNPAAWVRP 640
Db 601 SLQDQPTNLAAFPASHDPPAI PAPTATPSARGAVIRQNSDPTSEGPGSPNPAAWVRP 660
Qy 641 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRPRSNASAWQIYLORRAERGTPKPPG 700
Db 661 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRA----- 695
Qy 701 PPAQPPGPNASNDLRRSDPGWERSDVL PASHGLPQAGSLERNRVGASSKLDSSPV 760
Db 696 -----SNPDLRRSDPGWERSDVL PASHGLPQAGSLERNRVGSSKPDSSPV 743
Qy 761 LSPGNKAPDDHRSRPGPADPVLKERTLDEAPPPKAMDYSSSSEVESSEDEEG 820
Db 744 LSPGNKAPDDHRSRPGPADPVLKERTLDEAPPPKAMDYSSSSEVESSEDEEG 803
Qy 821 EGGPAEGSRDTPGGRSDGDTDSVSTWVHDVBEITGTPPYGGGTWVQRTPEERNLLH 880
Db 804 EGGPAEGSRDTPGGRSDGDTDSVSTWVHDVBEITGTPPYGGGTWVQRTPEERNLLH 863
Qy 881 ADSNGYTNLPDVVQPSHSPTENSKGSPSKDGDYQSRGLVAPKXSSFTMFVDLGIY 940
Db 864 ADSNGYTNLPDVVQPSHSPTENSKGSPSKDGDYQSRGLVAPKXSSFTMFVDLGIY 923
Qy 941 QPGSGSDSIPITALVGGEGTRLDQLOYDVYRKGSVNVNPTNTRAHSETPEIRKYKRFNS 1000
Db 924 QPGSGSDSIPITALVGGEGTRLDQLOYDVYRKGSVNVNPTNTRAHSETPEIRKYKRFNS 983
Qy 1001 ETLCAALMGVNLVGTENGLMLDRSGGKGYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
Db 984 ETLCAALMGVNLVGTENGLMLDRSGGKGYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
Qy 1061 KLRVYLSWLRNKLHNDPEVEKKQGTWTVGDMGCGHYRVVYKRIKFLVIALKSSVEV 1120
Db 1044 KLRVYLSWLRNKLHNDPEVEKKQGTWTVGDMGCGHYRVVYKRIKFLVIALKSSVEV 1103
Qy 1121 YAWAPKPHKFAFKSPADLPHRPLLDLTVEEGORLKVYIGSSAGFAHVDVDSGNSYDI 1180
Db 1104 YAWAPKPHKFAFKSPADLPHRPLLDLTVEEGORLKVYIGSSAGFAHVDVDSGNSYDI 1163
Qy 1181 YIPVHIQSOITPHALIFLPTDGMELLCYDEGVVYNTYGR I KDVVLQWGMPTSVAY 1240
Db 1164 YIPVHIQSOITPHALIFLPTDGMELLCYDEGVVYNTYGR I KDVVLQWGMPTSVAY 1223
Qy 1241 ICNQIMGWGEKAIERSVETGHLDGVMFKQAORLKFCLERNKDVFFASVRSGSSQVY 1300
Db 1224 ICNQIMGWGEKAIERSVETGHLDGVMFKQAORLKFCLERNKDVFFASVRSGSSQVY 1283
Qy 1301 FMTLNRNCIMNW 1312
Db 1284 FMTLNRNCIMNW 1295

RESULT 3

US-09-789-390-32
; Sequence 32, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-32

Query Match 96.1%; Score 6660.5; DB 10; Length 1295;
Best Local Similarity 95.6%; Pred. No. 1e-298;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

Qy 1 MGDPAPARSLDDIIDLALRDPAGIFELVVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Db 1 MGDPAPARSLDDIIDLALRDPAGIFELVVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Qy 61 DEEBEIKQEIINMLKKYSHHRIATYYGAFIKKSPGNDQDLWLVMEFCGAGSVTDLVKNT 120
Db 61 DEEBEIKQEIINMLKKYSHHRIATYYGAFIKKSPGNDQDLWLVMEFCGAGSVTDLVKNT 120
Qy 121 KGNALKECDIAYICREILRGLAHLHAKVTHRDIKGNVLLTENAELVLDVFGVSAQLDR 180
Db 121 KGNALKECDIAYICREILRGLAHLHAKVTHRDIKGNVLLTENAELVLDVFGVSAQLDR 180
Qy 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDINSGLITATEMAEGAPPLCDMHPMR 240
Db 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDINSGLITATEMAEGAPPLCDMHPMR 240
Qy 241 ALFLIPRNPPLKSKKSKKFIIDTCLITKYLSPPTQLLKFPPIRQPTQVR I 300
Db 241 ALFLIPRNPPLKSKKSKKFIIDTCLITKYLSPPTQLLKFPPIRQPTQVR I 300
Qy 301 QLKDHIDRSRKRGEKETETEYSGSEEDDSHGECEGEPSSIMNVPGESTLRRERFLRQQ 360
Db 301 QLKDHIDRSRKRGEKETETEYSGSEEDDSHGECEGEPSSIMNVPGESTLRRERFLRQQ 360
Qy 361 ENKNSBALKQOQLOKQOQOQDPEAHIKHLHQRRRIEEOKEERRRVEEQRREREOR 420
Db 361 ENKNSBALKQOQLOKQOQOQDPEAHIKHLHQRRRIEEOKEERRRVEEQRREREOR 420
Qy 421 LQEKQORRLDMOALREERROAEREQYKRLQLEEQRQSERLQRLQOQEHAYLSLQ 480
Db 421 LQEKQORRLDMOALREERROAEREQYKRLQLEEQRQSERLQRLQOQEHAYLSLQ 480
Qy 481 QOQOQOQLOKQOQOQOQLLPGRKPLHYHGRGMNPKADPAWAVEVERTRMKNQKNSPLAKS 540
Db 481 QOQOQOQLOKQOQOQOQLLPGRKPLHYHGRGMNPKADPAWAVEVERTRMKNQKNSPLAKS 540
Qy 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMORPVEPQEGPHKSLVAHRVPLKPYAAPVPRSQ 600
Db 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMORPVEPQEGPHKSLVAHRVPLKPYAAPVPRSQ 600
Qy 581 SLQDQPTNLAAFPASHDPPAI PAPTATPSARGAVIRQNSDPTSEGPGSPNPAAWVRP 640
Db 601 SLQDQPTNLAAFPASHDPPAI PAPTATPSARGAVIRQNSDPTSEGPGSPNPAAWVRP 660
Qy 641 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRPRSNASAWQIYLORRAERGTPKPPG 700
Db 661 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRA----- 695
Qy 701 PPAQPPGPNASNDLRRSDPGWERSDVL PASHGLPQAGSLERNRVGASSKLDSSPV 760
Db 696 -----SNPDLRRSDPGWERSDVL PASHGLPQAGSLERNRVGSSKPDSSPV 743


```
QY 1001 EILCAALWGNNLVGTENGMLLDRSGGQGVYGLIGRRRPOQMDVLEGLNLLITISGKRN 1060
DB 984 EILCAALWGNNLVGTENGMLLDRSGGQGVYGLIGRRRPOQMDVLEGLNLLITISGKRN 1043
QY 1061 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGEGCHYRVVVKYERIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGEGCHYRVVVKYERIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDI 1180
DB 1104 YAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDI 1163
QY 1181 YIPVHIQSQTTHAIIFLPTDGMELLCYEDEGVYVNTYGRIIKDVVLQWEMPTSVAY 1240
DB 1164 YIPVHIQSQTTHAIIFLPTDGMELLCYEDEGVYVNTYGRIIKDVVLQWEMPTSVAY 1223
QY 1241 ICSNQIMGWGEKAIIRSVETGHLQGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1300
DB 1224 ICSNQIMGWGEKAIIRSVETGHLQGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1283
QY 1301 FMTLNRNCIMNW 1312
DB 1284 FMTLNRNCIMNW 1295
```

RESULT 5

```
US-09-789-390-37
; Sequence 37, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-37
```

```
Query Match 96.1%; Score 6660.5; DB 10; Length 1295;
Best Local Similarity 95.6%; Pred. No. 1e-298;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

QY 1 MGDPAARSLODDIDLSALRDPAGIFELVEVVGNGTYGYQYGRHVKYTGQLAAIKVMDYTE 60
DB 1 MGDPAARSLODDIDLSALRDPAGIFELVEVVGNGTYGYQYGRHVKYTGQLAAIKVMDYTE 60
QY 61 DEEEIKQEIINMLKKYSHRNATATYGAFIKSPGNDQDLWLVMFCGAGSVTLVKNT 120
DB 61 DEEEIKQEIINMLKKYSHRNATATYGAFIKSPGNDQDLWLVMFCGAGSVTLVKNT 120
QY 121 KGNALKEDCIAYICREILRGLAHLHAHKVHHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHKVHHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
```

```
QY 181 TVGRNTEIGTYWMAPEVIACDENPDATYDYSIDIMSLGITALEMAGAPPLCDMHPMR 240
DB 181 TVGRNTEIGTYWMAPEVIACDENPDATYDYSIDIMSLGITALEMAGAPPLCDMHPMR 240
QY 241 ALFLIPRNPPLKSKWKKFIDFIDTCLTKTYLSRPPTEQLLKFPPIRQOPTERQVRI 300
DB 241 ALFLIPRNPPLKSKWKKFIDFIDTCLTKTYLSRPPTEQLLKFPPIRQOPTERQVRI 300
QY 301 QLKDHIDRSRKGRKEKETETEYSGSEEDDSHGEEGEPSSIMNVPGSESTLRRREFLRLOQ 360
DB 301 QLKDHIDRSRKGRKEKETETEYSGSEEDDSHGEEGEPSSIMNVPGSESTLRRREFLRLOQ 360
QY 361 ENKNSSEALKQOQOQLQOQOQORDPEAHIKHLHLHQRRLEEKQEKERRRVEEQORRERORK 420
DB 361 ENKNSSEALKQOQOQLQOQOQORDPEAHIKHLHLHQRRLEEKQEKERRRVEEQORRERORK 420
QY 421 LQEKQERRLQEDMOALRREERROAEREQYKRLQLEEQORQSERLQRLQOQEHAYLKSILQ 480
DB 421 LQEKQERRLQEDMOALRREERROAEREQYKRLQLEEQORQSERLQRLQOQEHAYLKSILQ 480
QY 481 QOQOQOQLQOQOQOQLLPQDRKPLYHYGRGNVPADKPAWAREVEBERTMKNQONSPLAKS 540
DB 481 QOQOQOQLQOQOQOQLLPQDRKPLYHYGRGNVPADKPAWAREVEBERTMKNQONSPLAKS 540
QY 541 KPGSTGPEPPIPOASPGPGPLSOTPPMORPVEPOEGPHK----- 580
DB 541 KPGSTGPEPPIPOASPGPGPLSOTPPMORPVEPOEGPHKSLVAHRVLPKYAAPVPRSQ 600
QY 581 SLQOQPTNLAAFPASHDPDPAPATPATPSARGAVIRQNSDPTSEGGPGSPNPAPWVRP 640
DB 601 SLQOQPTNLAAFPASHDPDPAPATPATPSARGAVIRQNSDPTSEGGPGSPNPAPWVRP 660
QY 641 DNEAPPKVPQRTSIIATALNTSGAGGRPAQAQVRA----- 700
DB 661 DNEAPPKVPQRTSIIATALNTSGAGGRPAQAQVRA----- 695
QY 701 PPAQPPGPNASSNPDLRRSDPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSPV 760
DB 696 -----SNPDLRRSDPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSPV 743
QY 761 LSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSSSEVESSEDEDEEG 820
DB 744 LSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSSSEVESSEDEDEEG 803
QY 821 EGGPAEGSRDTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGGTVMVQRTPEBERNLLH 880
DB 804 EGGPAEGSRDTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGGTVMVQRTPEBERNLLH 863
QY 881 ADSNGYTNLPDVVQPSHSPTESSKQSPSGDGYQSRGLVKAPGKSSFTMFVDLGIY 940
DB 864 ADSNGYTNLPDVVQPSHSPTESSKQSPSGDGYQSRGLVKAPGKSSFTMFVDLGIY 923
QY 941 QPGSGSDSIPITALVGGEGTRLDQLOQYDVRKGSVVNVNPTVTRAHSETPEIRKYKCRFNS 1000
DB 924 QPGSGSDSIPITALVGGEGTRLDQLOQYDVRKGSVVNVNPTVTRAHSETPEIRKYKCRFNS 983
QY 1001 EILCAALWGNNLVGTENGMLLDRSGGQGVYGLIGRRRFOQMDVLEGLNLLITISGKRN 1060
DB 984 EILCAALWGNNLVGTENGMLLDRSGGQGVYGLIGRRRFOQMDVLEGLNLLITISGKRN 1043
QY 1061 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGEGCHYRVVVKYERIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGEGCHYRVVVKYERIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEQRLKVIYSSAGFAHVDVDSGNSYDI 1180
DB 1104 YAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEQRLKVIYSSAGFAHVDVDSGNSYDI 1163
QY 1181 YIPVHIQSQTTHAIIFLPTDGMELLCYEDEGVYVNTYGRIIKDVVLQWEMPTSVAY 1240
DB 1164 YIPVHIQSQTTHAIIFLPTDGMELLCYEDEGVYVNTYGRIIKDVVLQWEMPTSVAY 1223
QY 1241 ICSNQIMGWGEKAIIRSVETGHLQGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1300
```

```

Db      1224  ICSNQIMGWGEKAIEIRSVETGHLGDFVFMHKAQRLKFLCERNDKVFFASVRSQSSQVY 1283
QY      1301  FMTLNRNCIMNW 1312
Db      1284  FMTLNRNCIMNW 1295

RESULT 6
US-09-789-390-39
; Sequence 39, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-39

Query Match      96.1%; Score 6660.5; DB 10; Length 1295;
Best Local Similarity 95.6%; Pred. No. 1e-298;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

QY      1  MGDPAARSDDIDLSALADPDAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Db      1  MGDPAARSDDIDLSALADPDAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
QY      61  DEEBEIKQBINMLKXYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTDLVKNT 120
Db      61  DEEBEIKQBINMLKXYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTDLVKNT 120
QY      121  KGNALKEDCIAYICREILGLAHLAHKVIHDDIKQNVLLTENAEEVKLVDFGVSQAQLDR 180
Db      121  KGNALKEDCIAYICREILGLAHLAHKVIHDDIKQNVLLTENAEEVKLVDFGVSQAQLDR 180
QY      181  TVGRNRTFTGTYPWNAPEVIACDENPDATYDYSRDIWSLIGITAIEMAGCAPLCDHMPWR 240
Db      181  TVGRNRTFTGTYPWNAPEVIACDENPDATYDYSRDIWSLIGITAIEMAGCAPLCDHMPWR 240
QY      241  ALFLIPRNPPLKSKKSKKFIIDFTCLIKTYLSRPPTTEQLLKFPPIRDOPTERQVRI 300
Db      241  ALFLIPRNPPLKSKKSKKFIIDFTCLIKTYLSRPPTTEQLLKFPPIRDOPTERQVRI 300
QY      301  QLKDHIDRSRKRGEKEETEYVSGSEEDDSSHGEPSIMNVPGESTLRREPLRLOQ 360
Db      301  QLKDHIDRSRKRGEKEETEYVSGSEEDDSSHGEPSIMNVPGESTLRREPLRLOQ 360
QY      361  ENKNSSEALKQQQQLKQQQQRDPEAHIKHLHQRRRIIEEQKEERRRVEEQRREREQRK 420
Db      361  ENKNSSEALKQQQQLKQQQQRDPEAHIKHLHQRRRIIEEQKEERRRVEEQRREREQRK 420
QY      421  LOEKSQRRLEDMQALRRBEERRQAEREOQYKRLQLEBQROSERLQRLQOEHA YKLSLQ 480

```

```

Db      421  LQEKEQRRLEDMQALRRBEERRQAEREOQYKRLQLEBQROSERLQRLQOEHA YKLSLQ 480
QY      481  QQQQQQQLKQQQQLKQQQQRDPEAHIKHLHQRRRIIEEQKEERRRVEEQRREREQRK 540
Db      481  QQQQQQQLKQQQQLKQQQQRDPEAHIKHLHQRRRIIEEQKEERRRVEEQRREREQRK 540
QY      541  KPGSTGPEPPIPOASPGPPGPLSQTTPMQRVPEQEGPHK----- 580
Db      541  KPGSTGPEPPIPOASPGPPGPLSQTTPMQRVPEQEGPHKSLVAHRVPLKYAAPVPSRQ 600
QY      581  SLQDQPTRNLAAFPASHDDPDPAIPAPTATPSARGAVIRONSDPTSESGPGSPNPAPWVRP 640
Db      601  SLQDQPTRNLAAFPASHDDPDPAIPAPTATPSARGAVIRONSDPTSESGPGSPNPAPWVRP 660
QY      641  DNEAPPKVPORSTSIATALTSGAGSRPAQAVARPRNSAWQIYLQRBARTGTPKPPG 700
Db      661  DNEAPPKVPORSTSIATALTSGAGSRPAQAVARA----- 695
QY      701  PPAQPPGPPPNASNPDLRRSDPGWERSDVLPSHGHLPQAGSLERNRVGASSKLDSSPV 760
Db      696  -----SNPDLRRSDPGWERSDVLPSHGHLPQAGSLERNRVGSSKXPDSPV 743
QY      761  LSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSESEVEDDEBEG 820
Db      744  LSPGNKAKPDHRSRGRPADFVLLKERTLDEAPRPPKAMDYSSSESEVEDDEBEG 803
QY      821  EGGPAEGSRDTPGGRSDGTDTSVSTMVVDVVEETGTQPPYGGGTVMVQVTPPEERNLLH 880
Db      804  EGGPAEGSRDTPGGRSDGTDTSVSTMVVDVVEETGTQPPYGGGTVMVQVTPPEERNLLH 863
QY      881  ADSNGYTNLPDVVQPSHSPTESSKQSPSPKSGDYGQSRGLVKAPGKSSFTMFVLDGIY 940
Db      864  ADSNGYTNLPDVVQPSHSPTESSKQSPSPKSGDYGQSRGLVKAPGKSSFTMFVLDGIY 923
QY      941  QPGSGSDSIPITALVGGEGTRLQDYDVRKGSVVNNVNTNTRAHSETPEIRKYKGRFNS 1000
Db      924  QPGSGSDSIPITALVGGEGTRLQDYDVRKGSVVNNVNTNTRAHSETPEIRKYKGRFNS 983
QY      1001  EILCAALGWVNLVGTENGLMLDRSGOGKYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
Db      984  EILCAALGWVNLVGTENGLMLDRSGOGKYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
QY      1061  KLRVYLSLWLNKILHNDPEVEKKQGTWTYGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1120
Db      1044  KLRVYLSLWLNKILHNDPEVEKKQGTWTYGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1103
QY      1121  YAWAPKPYHKFMAFKSFADLPHRPLLDLTVEEQRLKVIYSSAGFAHVDVDSGNSYDI 1180
Db      1104  YAWAPKPYHKFMAFKSFADLPHRPLLDLTVEEQRLKVIYSSAGFAHVDVDSGNSYDI 1163
QY      1181  YIPVHIQSOITPHAIIFLPNTDGMELLCYDEDEGVYVNTYGRITKDVVLQWGMPTSVAY 1240
Db      1164  YIPVHIQSOITPHAIIFLPNTDGMELLCYDEDEGVYVNTYGRITKDVVLQWGMPTSVAY 1223
QY      1241  ICSNQIMGWGEKAIEIRSVETGHLGDFVFMHKAQRLKFLCERNDKVFFASVRSQSSQVY 1300
Db      1224  ICSNQIMGWGEKAIEIRSVETGHLGDFVFMHKAQRLKFLCERNDKVFFASVRSQSSQVY 1283
QY      1301  FMTLNRNCIMNW 1312
Db      1284  FMTLNRNCIMNW 1295

RESULT 7
US-09-291-417-15
; Sequence 15, Application US/09291417A
; Publication No. US20030050230A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

```


Db 241 ALFLIPRPPPLKSKKKKKKFIIDTCLITKYLISRPPTQLLAPFPFIRDQPTQVR 300
QY 301 QLKDHIDRSKRGKEETEYISGSEEDDGHGEGPSSIMNVPGESTLRRFLRLQQ 360
Db 301 QLKDHIDRSKRGKEETEYISGSEEDDGHGEGPSSIMNVPGESTLRRFLRLQQ 360
QY 361 ENKNSSEALKQQQQLQQQQORPEAHIKHLLHQRORRIIEQKEERRRVEEQORREORK 420
Db 361 ENKNSSEALKQQQQLQQQQORPEAHIKHLLHQRORRIIEQKEERRRVEEQORREORK 420
QY 421 LOEKQORLEDMQALREERERRQAEQYKRLKLEQORQSERLQRLQOQHAYLKSQ 480
Db 421 LOEKQORLEDMQALREERERRQAEQYKRLKLEQORQSERLQRLQOQHAYLKSQ 480
QY 481 QQQQQQQQLKQQQQQLLPGDKRPLHYHGRGNMADKPAWAREVEERTMKNQNSPLAKS 540
Db 481 QQQQQQQQLKQQQQQLLPGDKRPLHYHGRGNMADKPAWAREVEERTMKNQNSPLAKS 540
QY 541 KPGSTGPEPPIQASGPGPGPLSQTPMQRVPEQEGPHK----- 580
Db 541 KPGSTGPEPPIQASGPGPGPLSQTPMQRVPEQEGPHKLSVAHRVPLKPYAAPVPSRQ 600
QY 581 SIQDQPTRNLAAPFASHDDPAIPAPTATPSARGAVIRQNSDPTSGPGSPNPAAWRP 640
Db 601 SIQDQPTRNLAAPFASHDDPAIPAPTATPSARGAVIRQNSDPTSGPGSPNPAAWRP 660
QY 641 DNEAPPKVPQRTSSITATLNTSGAGSRPAQAVRAPRPSNSAWQIYLQRRRAERGTPPKPG 700
Db 661 DNEAPPKVPQRTSSITATLNTSGAGSRPAQAVRA----- 695
QY 701 PPAQPGPNASSNPDLRRSDPGWRSDSVLPASHGHLPAQGLSERNRVGASSKLDSPV 760
Db 696 -----SNPDLRRSDPGWRSDSVLPASHGHLPAQGLSERNRVGSSKPDSPV 743
QY 761 LSPGNKAPDDHRSRPGPA-----DFVLLKERTLDEAPRPKKAWDYSSSSEEVES 812
Db 744 LSPGNKAPDDHRSRPGRPASVKRAIGEDFVLLKERTLDEAPRPKKAWDYSSSSEEVES 803
QY 813 SEDDEEGEGGPAEGSRDTPGCRSDGSDSVSTMVHVHVEITGTQPPYGGTMMVQRT 872
Db 804 SEDDEEGEGGPAEGSRDTPGCRSDGSDSVSTMVHVHVEITGTQPPYGGTMMVQRT 863
QY 873 BEERNLLHADSNGYTNLDDVVQPSHSPTENSKGQSPSKDGGDVQSRGLVAPKQSSFT 932
Db 864 BEERNLLHADSNGYTNLDDVVQPSHSPTENSKGQSPSKDGGDVQSRGLVAPKQSSFT 923
QY 933 MPVDLGIYQPGSGDSIPTALVGEGTRLQOYDVRKGSVYVNPNTTRAHSETPEIR 992
Db 924 MPVDLGIYQPGSGDSIPTALVGEGTRLQOYDVRKGSVYVNPNTTRAHSETPEIR 983
QY 993 KYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGVYGLIGRRRFPQMDVLEGLNL 1052
Db 984 KYKRFNSEILCAALWGVNLLVGTENGMLLDRSGQGVYGLIGRRRFPQMDVLEGLNL 1043
QY 1053 ITISGRNKLRYVYLSWLNKILHNDPEVEKKQGTMTVDMEGCGHYRVKYERIKFLVI 1112
Db 1044 ITISGRNKLRYVYLSWLNKILHNDPEVEKKQGTMTVDMEGCGHYRVKYERIKFLVI 1103
QY 1113 ALKSSVEVYAWAPKBYHKFMAFKSADLPHRPLLDLTVREGQRLKVIYSSAGHAYDV 1172
Db 1104 ALKSSVEVYAWAPKBYHKFMAFKSADLPHRPLLDLTVREGQRLKVIYSSAGHAYDV 1163
QY 1173 DSGNSYDIYIPVHIQSQTPIHAIIFLPTNDGEMLLCYEDEGVYNTYGRIIKDVVLQW 1232
Db 1164 DSGNSYDIYIPVHIQSQTPIHAIIFLPTNDGEMLLCYEDEGVYNTYGRIIKDVVLQW 1223
QY 1233 EMPTSVAYICSNQIMGWGEKAIERSVETGHLDGVPFMHKAQRLKFLCERNDKVFFASVR 1292
Db 1224 EMPTSVAYICSNQIMGWGEKAIERSVETGHLDGVPFMHKAQRLKFLCERNDKVFFASVR 1283
QY 1293 SGGSSQVYFMTLNRNCIMNW 1312

Db 1284 SGGSSQVYFMTLNRNCIMNW 1303
RESULT 10
US-09-789-390-9
; Sequence 9, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shimkous, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-9

Query Match 95.7%; Score 6628.5; DB 10; Length 1303;
Best Local Similarity 94.8%; Pred. No. 3.1e-297; Mismatches 1; Indels 65; Gaps 3;
Matches 1270; Conservative 1;

QY 1 MGDPAPARSLDDIDLSALRDPAGIFELVEVVGNGTYGQVYKGRHVKTGQLAAIKVMDYTE 60
Db 1 MGDPAPARSLDDIDLSALRDPAGIFELVEVVGNGTYGQVYKGRHVKTGQLAAIKVMDYTE 60
QY 61 DEEEIKQEINMLKKYSHHRNIATYYGAFIKKSPGNDQDLWLMFCGAGSVTDLVKNT 120
Db 61 DEEEIKQEINMLKKYSHHRNIATYYGAFIKKSPGNDQDLWLMFCGAGSVTDLVKNT 120
QY 121 KGNALKEDCTAYICREILRGLAHHAHKVTHRDIKGNVLLTENAENVKLVDFGVSAQLDR 180
Db 121 KGNALKEDCTAYICREILRGLAHHAHKVTHRDIKGNVLLTENAENVKLVDFGVSAQLDR 180
QY 181 TVGRNRTFICTPYWMAPEVIACDENPDATYDYSRDIWSLGITAIEMAEGAPPLCDMHPMR 240
Db 181 TVGRNRTFICTPYWMAPEVIACDENPDATYDYSRDIWSLGITAIEMAEGAPPLCDMHPMR 240
QY 241 ALFLIPNPPRLKSKKSKKFIIDTCLITKYLSPPTQLLAPFPFIRDQPTQVR 300
Db 241 ALFLIPNPPRLKSKKSKKFIIDTCLITKYLSPPTQLLAPFPFIRDQPTQVR 300
QY 301 QLKDHIDRSKRGKEETEYISGSEEDDGHGEGPSSIMNVPGESTLRRFLRLQQ 360
Db 301 QLKDHIDRSKRGKEETEYISGSEEDDGHGEGPSSIMNVPGESTLRRFLRLQQ 360
QY 361 ENKNSSEALKQQQQLQQQQORPEAHIKHLLHQRORRIIEQKEERRRVEEQORREORK 420
Db 361 ENKNSSEALKQQQQLQQQQORPEAHIKHLLHQRORRIIEQKEERRRVEEQORREORK 420
QY 421 LOEKQORLEDMQALREERERRQAEQYKRLKLEQORQSERLQRLQOQHAYLKSQ 480
Db 421 LOEKQORLEDMQALREERERRQAEQYKRLKLEQORQSERLQRLQOQHAYLKSQ 480
QY 481 QQQQQQQQLKQQQQQLLPGDKRPLHYHGRGNMADKPAWAREVEERTMKNQNSPLAKS 540


```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(1303)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification.
US-09-789-390-65

Query Match      78.3%; Score 5422.5; DB 10; Length 1303;
Best Local Similarity 79.9%; Pred. No. 11e-241;
Matches 1071; Conservative 0; Mismatches 204; Indels 65; Gaps 3;

Qy 1 MGDPAARSLLDIDLSALRDPAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Db 1 MGDPAARSLLDIDLSALRDPAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Qy 61 DEEBEIKQBINMLKYSKSHRNATYTGAPIKSPGNDQDLWLMVEFCAGSVTLVKNT 120
Db 61 XXXXXIKQBINMLKYSKSHRNATYTGAPIKSPGNDQDLWLMVEFCAGSVTLVKNT 120
Qy 121 KGNALKEDCICATYICREILGLAHLHAHKVHRIKQNVLLTENAENVKLVDFGSAQLDR 180
Db 121 KGNALKEDCICATYICREILGLAHLHAHKVHRIKQNVLLTENAENVKLVDFGSAQLDR 180
Qy 181 TVGRRNTFTGTPYMAPEVIAIDENPDATYDYSRDIWSLGITAIEMAGAPPLCDMHPWR 240
Db 181 TVGRRNTFTGTPYMAPEVIAIDENPDATYDYSRDIWSLGITAIEMAGAPPLCDMHPWR 240
Qy 241 ALFLIPRNPPLKSKGSKKIFIDTCLIKTYLSRPPTQQLKFPFIDQPTQVRRI 300
Db 241 ALFLIPRNPPLKSKGSKKIFIDTCLIKTYLSRPPTQQLKFPFIDQPTQVRRI 300
Qy 301 QLKDHIDRSRKRKGEKEEYEVYSGSEEDDSDGEGEPSSIMVPGESTLRREFLRQQ 360
Db 301 QLKDHIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXPSIMVPGESTLRREFLRQQ 360
Qy 361 ENKNSSEALKOQQOQLQOQQOORDEPAHIKHLHQRRIEQEKEERRRVEEQORREREQRK 420
Db 361 ENKNSSEALKXXXXXXXXXXRRDPEAHIKHLHXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 420
Qy 421 LOEKQOORLEDMQALRREERERRQREOEYKRLQLEQRQSERLQRLQOEHAYLSLQ 480
Db 421 XXXXXXXXXXXXXDMQALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKSKX 480
Qy 481 QOQQOQQOQLQOQQOQLLPGDKRPLHYGRGMNPAKPAWAREVEERTMKNQONSPLAKS 540
Db 481 XXXXXXXXXXXXXXXXXXXXXPGDKRPLHYGRGMNPAKPAWAREVEERTMKNQONSPLAKS 540
Qy 541 KPGSTGPEPPIPOASPGPPGLSQTTPMQRPVPEQEGPHK----- 580
Db 541 KPGSTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXMQRPVPEQEGPHKSLVAHRVPLKPYAAPVPRSQ 600
Qy 581 SLQDQPTRNLAAFPASHDDPAIPAPTATPSARGAVIRQNSDPTSEGGPGSPNPAPWVRP 640
Db 601 SLQDQPTRNLAAFPASHXXXXXXXXXXXXXXXXXXXXXRGAVIRQNSDPTSEGGPGSPNPAPWVRP 660
Qy 641 DNEAPPKVPQRTSSITATANTSGAGSRPAQAVRARPRNSAWQIYLQRAERGTPKPPG 700
Db 661 DNEAPPKVPQRTSSITATANTSGAGSRPAQAVRA----- 695
Qy 701 PPAQPPGPNASNPDLRRSDPGWERSDVLPAASHGLHPQAGSLERNRVGASSKLDSPV 760
Db 696 -----SNPDLRRSDPGWERSDVLPAASHGLHPQAGSLERNRVGSSKPDSPV 743
Qy 761 LSPGNKAKPDDHRSRGRPA-----DFVLLKERTLDEAPRPPKAKMDYSSSEEVES 812
Db 744 LSPGNKAKPDDHRSRGRPASYPASYKRAIGEDFVLLKERTLDEAPRPPKAKMDYXXXXXXXXXX 803
Qy 813 SEDDDEEGEGGPAEGSRDTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGGTWVQRTTP 872
Db 804 XXXXXXXXXXXXXXXXRRDTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGGTWVQRTTP 863
```

Search completed: August 28, 2004, 01:03:28
Job time : 88 sec